

GENE CLUSTER FOR PRODUCTION OF THE ENEDIYNE
ANTITUMOR ANTIBIOTIC C-1027

SEQUENCE LISTING

SEQ ID No. 1. C-1027 gene cluster DNA sequence from 1 to 42,000, ORF-(-7) to ORF-26

1 GTCGACTCTAGAGGATCCCGGGTGCAGGAGTAGGGGTTACGGACGAAGGAGGGGTGCCCGG
-----+-----+-----+-----+-----+ 60
-7-* CAGCTGAGATCTCCTAGGGCCACGCCTCATCCCCAATGCCTGCTTCTCCACGGGCC
* L I G P A S Y P N R V F S P H G -
CGACGCCTGCGGCGAAGGGCGGTTCTTCTGAGTTCGAGGCCGGTGGCGAGGACGACGTGGT
61 -----+-----+-----+-----+-----+ 120
-7 GCTGCGGACGCCGCTTCCCGCAAGGAATCAAGTCCCGGCCACCGCTCCTGCTGCACCA
A V G A A F P P E K L E L G T A L V V H -
CCGCGTCGAGGATCTGCGTGTGCGGGAGCGGCCAGGGCGCAGCCCCTCGGTGAGGTACG
121 -----+-----+-----+-----+-----+ 180
-7 GGCGCAGCTCCTAGACGCACAGCCCCTCGCGGGTCCCGCGTCGCGGAGCCAGTCCATGC
D A D L I Q T D P L P G P R L G E T L Y -
GGGTGAGGCCCTGACGGTCACCTCGAAGCAGCGGTCGTGGGACCGGGCGTCGAGCGCCT
181 -----+-----+-----+-----+-----+ 240
-7 CCCACTCCGGGGACTGCCAGTGGAGCTTCGTGCGCCAGCACCTGGCCCGCAGCTCGCGGA
P T L G R V T V E F C R D H S R A D L A -
CCCCGTCCGCTTCCACAAGGACGACCGCGGGACAGGACTCCCGTGGCGCCTCGACCAGTC
241 -----+-----+-----+-----+-----+ 300
-7 GGGGCAGGCGAAGGTGTCTGCTGCGGCCCTGCTCTGAGGGCACGCCGAGCTGGTCAG
E G D A E V L V V G P C S E R A A E V L -
GGGCGTCGAGGTAGTCTTGAAGATGCGGCGGGGGCGGGGCCCTGTTGCGTGAAGTCC
301 -----+-----+-----+-----+-----+ 360
-7 CCCGCAGCTCCATCAGGACCTTCTACGCCGCCCGCGCCCGGACCAAGCCACTTGAAGG
R A D L Y D Q F I R R P A P G Q E T F K -
ACGAAGCCCAGCGCCGGGGCCAGTCGCGCCGGTCCGCTCCTGGTTGGCCAGTTGATGA
361 -----+-----+-----+-----+-----+ 420
-7 TGCTTCGGGTGCGGCCCCGGTCAGCGCGGCCAGCGGAGGACCAACCGGGTCAACTACT
W S A W R R P W D R R D A E Q N A W N I -
AGTCGAGCACGTCTCGCGGAACACCGACATCCTGCGGCCTGGATATTGAAGACGTGGT
421 -----+-----+-----+-----+-----+ 480
-7 TCAGCTCGTGCAGGAGCGCCTTGTGGCTGTAGGACGCCCGGACCTATAACTTCTGCACCA
F D L V D E R F V S M R G A Q I N F V H -
CCCAGGGGTGCGGTCACGGTGATAGGCGACGCCGGCCGAGCGGTAGCGGCGCGCCGCT
481 -----+-----+-----+-----+-----+ 540
-7 GGGTCCCCAACGGCAGTGCCACTATCCGCTGCGGCGCGCTCGCCATCGCCGCGCGGCGA
D W P N G D R H Y A V G A S R Y A A R R -
CCAGGAGGACGACTTCCAGCGGTCTTCTCGCGAAATGAAGCAGGCGTATCGCGGTGCGCG
541 -----+-----+-----+-----+-----+ 600
-7 GGTCTCTGCTGAAGGTGCGCCAGAAGAGCGCTTTACTTCGTCCGCATAGCGCCAGCGGC
E L L V V E L P R R A F H L L R I A T A -
TGCCTGCCAGGCCCGCCCCTACGACCAGCACCTGGGGCGCGCACCCGTCATGCCCATGA
601 -----+-----+-----+-----+-----+ 660
-7-< ACGGACGGTCCGGGCGGGGATGCTGGTGTGGGACCCCGCGGTGGGCAGTACGGGTACT
T G A L G A G V V L V R P R A G T M G M -

AGCCTCCCCCGCTGACTCAGGGCGCGCGTCGCGCGCTCCCGTCGGTGTCTCGCTGACT
 661 -----+-----+-----+-----+-----+-----+ 720
 TCGGAGGGGGCGACTGAGTCCCGCgCGCGAGCGCGAGGGCAGCCACAGGAGCGACTGA

 GGAAGTTCCTGACCTGGCGTCAACTCCACTGATCCGTAAGGGGATCGCGGGAGTGGATA
 721 -----+-----+-----+-----+-----+-----+ 780
 CCTTCAAGGGACTGGACCGCAGTTGAGGTGACTAGGCATTCCCCTAGCGCCCTCACCTAT

 CGGGTCAGGTTCGTGCACGATCGTGGCACCAGACAGATCACCACGTGATAGGCACTCGTG
 781 -----+-----+-----+-----+-----+-----+ 840
 GCCCAGTCCAGCACGTGCTAGCACCGTGGTCTGTCTAGTGGTGCAGCTATCCGTGAGCAC

 AGCCGCGCCCGGGGCTCGACGGGGCGGGGCACCGGCAGGGGCGGCCGCGTGATCAGCCGG
 841 -----+-----+-----+-----+-----+-----+ 900
 TCGGCGCGGGCCCCGAGCTGCCCCGCCCCGTGGCCGTCCCCGCGGCGCACTAGTCGGCC

 AGCCTGTCCGGGGGCGTGCGTGCGGGGCGTCAGCTGTGATGTGCGGAACGCCAGGGACG
 901 -----+-----+-----+-----+-----+-----+ 960
 TCGGACAGGCCCCCGCACGCACGCCCCGAGTCGACAGCTACAGCCCTTGCGGTCCCTGC
 -6- * * S D I D P V G P V -

 TCGATCTCGGTGCGGGCGTAGTGGTTGAAGTAGTTGGTGTAGAGGTTACGGGCCACGTGG
 961 -----+-----+-----+-----+-----+-----+ 1020
 AGCTAGAGCCACGCCCCGCATCACCAACTTCATCAACCACATCTCCAAGTGCCGGTGCACC
 -6 D I E T R A Y H N F Y N T Y L N V A V H -

 ACGAAGACCTCGGCGAGCTCGGTGTCCGTCCATCCCTGTGCCACGGCCGCGTTCCACGAG
 1021 -----+-----+-----+-----+-----+-----+ 1080
 TGCTTCTGGAGCCGCTCGAGCCACAGGCAGGTAGGGACACGGTGCCGGCGCAAGGTGCTC
 -6 V F V E A L E T D T W G Q A V A A N W S -

 GCGTCAGACGCCTCGCCCACTTCGCCGGCGATCTCCCTGGCCACCTGGACCAGTGCTTCG
 1081 -----+-----+-----+-----+-----+-----+ 1140
 CGCAGTCTGCGGAGCGGGTGAAGCGGCCGCTAGAGGGACCGGTGGACCTGGTCACGAAGC
 -6 A D S A E G V E G A I E R A V Q V L A E -

 AGCTTCACGTTCGTGCGCGGGCGTCCCCCGGCAATCGCCACGGTCTCCTCCAGCGTGAAA
 1141 -----+-----+-----+-----+-----+-----+ 1200
 TCGAAGTGCAGCAGCGGCCCGCAGGGGGCCGCTTAGCGGTGCCAGAGGAGTTCGCACTTT
 -6 L K V D D G P T G R R I A V T E E L T F -

 CCCGCGACCTTCGCCGACACCGTGTGCGCCGCTGGCAGTACGCGCACGCGTCGACCGCG
 1201 -----+-----+-----+-----+-----+-----+ 1260
 GGGCGCTGGAAGCGGCTGTGGCACACGCGGCGGACCGTCATGCGCGTGCGCAGCTGGCGC
 -6 G A V K A S V T H A A Q C Y A C A D V A -

 CCCACGGCGAGGGCGATCGCCTCGCGTGTGCGGGCGTGAACGTTCCATGTTTCGGCGACG
 1261 -----+-----+-----+-----+-----+-----+ 1320
 GGGTGCCGCTCCCGCTAGCGGAGCGCACACGCCCCGAGCTTGCAAGGTACAAGCCGCTGC
 -6 G V A L A I A E R T R A D F T G H E A V -

 GCTCCGGTGATCGCGGCGTAGGTTTCCAGGACCACGGGGGAATGGGCCATTCCCCCGTGG
 1321 -----+-----+-----+-----+-----+-----+ 1380
 CGAGGCCACTAGCGCCGCATCAAAGGTCCTGGTGCCCCCTTACCCGGTAAGGGGGCACC
 -6 A G T I A A Y T E L V V P S H A M G G H -

 ATGTTGAGCACTCGCCCGAACCGCTTCTCCAGTCGGCGCAGGATGTCTCCGCCGGCTGCG
 1381 -----+-----+-----+-----+-----+-----+ 1440
 TACAACCTCGTGAGCGGGCTTGGCGAAGAGGTGAGCCGCGTCTACAGAGGCGGCCGACGC
 -6 I N L V R G F R K E L R R L I D G G A A -

 GGTGCGGTGTGATGGTGTGGACGGGAATCCGCGGCATGGGAATGCCTCTCCTCGTAGTG
 1441 -----+-----+-----+-----+-----+-----+ 1500
 CCACGCCACAGCTACCACACCAGCCCTTAGGCGCCGTACCCTTACGGAGAGGAGCATCAC
 -6- < P A T D I T H V P I R P M -

```

ATGGGAGTTCTCGTCCCTCCAGTCTGCCAAGCACCTCCCCCGGTGAGCTGTCCCGGCC
1501 -----+-----+-----+-----+-----+-----+ 1560
TACCTCAAGGAGCAGGGAGGTGACACGGGTTCTGAGAGGGGGCCACTCGACAGGGCCGG

GCCCTCCGGCCCCCTTCTAGGCAGGTGCCCCGGTGGTGCGGCCCCAGGACGTACCTCGCC
1561 -----+-----+-----+-----+-----+-----+ 1620
CGGGAGGCCGGGAAGATCCGTCCAGCGGGCCACCACGCCGGGGTCTGCACTGGAGCGG

GCACCACCGGGAGCCCCGAGGGGCGAGGTGAGAGGCCGAGCACCTCTCGGCCAGGGCGG
1621 -----+-----+-----+-----+-----+-----+ 1680
CGTGGTGGCCCTCGGGGCTCCCCGCTCCAGTCTCCGGCTCGTGGAGGAGCCGGTCCCCGCC
-5- * * L G L V E E A L A -

TGCCCCGAACACGGGCCTCGATCTTGGCGAAGGCCAGGTGCGGTGGTGGAGGTGTCTGT
1681 -----+-----+-----+-----+-----+-----+ 1740
ACGGGGCTTGTGCCCGGAGCTAGAACCCTTCCGGTCCAGCGCACACCACCTCCACAGCA
-5 T G R V R A E I K A F A L D R T T S T D -

CGCGAACGGGGAGAAGCCGAGTCTGCGAGGTTCCAGTTGCTCGACGGGGATGTAGC
1741 -----+-----+-----+-----+-----+-----+ 1800
GCCGCTTGGCCCTCTTCGGCGTCAAGAGGTTCAACGAGCTGCCCTACATCG
-5 D A F P S F G C D D C T G L Q E V P I Y -

GGGCGGCGAGCAGGATGCGGTGCGGTACCTGCTCGGGGGTCTCGACCACTGGGTGATCG
1801 -----+-----+-----+-----+-----+-----+ 1860
CCCGCGCTCGTCTACGCCAGCGCATGGACGAGCCCCAGAGCTGGTGACCCAGCTAGC
-5 R A A L L I R D R V Q E P T E V V P D I -

GGTCTGGTCACCCCGAGGAAGACGCGGGCGGCAGGGGGCAGGTGGTCACGGACGATGCTCA
1861 -----+-----+-----+-----+-----+-----+ 1920
CCAGCCAGTGGGGCTCCTTCTGCGCCCGCGTCCCCCGTCCACCACTGCTGCTACGAGT
-5 P D T V G L F V R A A P P L H D R V I S -

GGACCCGCTCGGGGTCCGCTTCGCGCGCCAGTTCGAGATAGAAGTTGCCCGCCTTGAGCT
1921 -----+-----+-----+-----+-----+-----+ 1980
CCTGGGCGAGCCCCAGGCGAAGCGGCGGTCAGCTCTATCTTCAACGGGCGGAAGTCA
-5 L V R E P D A E G A L E L Y F N G A K L -

GGAAGAGCTTGGGCGAGTTCGGCGTAGTCGATGTCGAGGCTGTGCGTGGAGTCTGGT
1981 -----+-----+-----+-----+-----+-----+ 2040
CCTTCTCGAACCCGTCGTCAAGCCGATCAGCTACAGCTCCGACACGCACCTCAGGACCA
-5 Q F L K P L L E A Y D I D L S H T S D Q -

CGCCGCCGGGGCAGGTGTGTACGCCGATGCGGGCGGTTTCTCGGCGCTGAAGCGCCCCA
2041 -----+-----+-----+-----+-----+-----+ 2100
GCGGCGGCCCCGTCCACACATGCGGCTACGCCCGCAAAGGAGCCGCGACTTCGCGGGGT
-5 D G G P C T H V G I R A T E E A S F R G -

GGAATTCGTTGTTGAGGGCGATGAAGTCGTCGAGGACGCCCGCTGGGGTTCGAGCTTGA
2101 -----+-----+-----+-----+-----+-----+ 2160
CCTGAAGCAACAACCTCCGCTACTTCAGCAGCTCCTGCGGCGGCGACCCAGCTCGAAGT
-5 L V E N N L A I F D D L V G G S P D L K -

GGGACAGCCGCCCTCGGTGAAGTCGAGCTGGACCACGTGTGCCCCCGCTCCAGGCAGC
2161 -----+-----+-----+-----+-----+-----+ 2220
CCCTGTCGGCGGGGAGCCACTTCAGCTCGACCTGGTGCACACGGGGGCGCAGGTCCGTCG
-5 L S L R G E T F D L Q V V H A G A D L C -

CTCGGATGTCGGCTTCGGCCTCGTCGGCGAGGTGCGCGCAGGAAGTCTCGCGGGGGTAGC
2221 -----+-----+-----+-----+-----+-----+ 2280
GAGCCTACAGCCGAAGCCGAGCAGCCGCTCCAGCGCTCCTTGACGAGCGCCCCCATCG
-5 G R I D A E A E D A L D R L F Q E R P Y -

CCTCGATGGGAGTGGCGGGGTAGAGGAGGTGAGGGCGGAGGGTGGCGATGACCGCTGCT
2281 -----+-----+-----+-----+-----+-----+ 2340

```

GGAGCTACCCTCACC GCCCCATCTCTCCGACTCCCGCCTCCCACGCTACTGGCGGACGA
 -5 G E I P T A P Y L L S L A S P A I V A Q -
 TCAGGGGGCGGTCCGTGAGCTGCCGTGCGGCGCGCAGATAGGTTTCGGCCCCGCACCTGGT
 2341 -----+-----+-----+-----+-----+ 2400
 AGTCCCCCGCCAGGCACTCGACGGCACGCCGCGCTCTATCCAAAGCCGGGCGTGGACCA
 -5 K L P R D T L Q R A A R L Y T E A R V Q -
 AGCGGAAGGGCCCTTGGGTGATGCTGGGGAGCTGCCGGGTGTGCCCGTCTGCGAAGGGGA
 2401 -----+-----+-----+-----+-----+ 2460
 TCGCCTTCCCCGGAACCCACTACGACCCCTCGACGGCCACACGGGCAGACGCTTCCCCCT
 -5 Y R F P G Q T I S P L Q R T H G D A F P -
 TGACAGCGCCGTGCGGCGAGAGGGTGTGAGGCCGTACGCGGGTAGGTGGCGAAGCTCG
 2461 -----+-----+-----+-----+-----+ 2520
 ACTGTGCGGGCAGCCCCGCTCTCCACAGCTCCGGCCAGTGCCCCATCCACCGCTTCGAGC
 -5 I V A G D P S L T D L G T V P Y T A F S -
 GCTTGGA CTGTTACCGTCCACGAGGACGGGGCTGCCGACTCGTTCCAGTCGTGTGAGGG
 2521 -----+-----+-----+-----+-----+ 2580
 CGAACCTGACAAGTGGCAGGTGCTCCTGCCCGACGGCTGAGCAAGGTGAGCACAGTCCC
 -5 P K S Q E G D V L V P S G V R E L R T L -
 TGTCCGCGACGGCCTGTTCTCTGCTGTTTGGCCAGGTCCGTGGCGTCCAGGGTTCCCTGGG
 2581 -----+-----+-----+-----+-----+ 2640
 ACAGGCGCTGCCGACAAGGACGACAAACCGGTCCAGGCACCGCAGGTCCCAAGGGACCC
 -5 T D A V A Q E Q Q K A L D T A D L T G Q -
 CATGCGCGGCAAGGGCGTGAGGAGTGTGCGGAGCGCGGAAGGCTGCCGATCGGCTCAG
 2641 -----+-----+-----+-----+-----+ 2700
 GTACGCGCCGTTCGCCGACGTCCTCACAGCGCCTCGCGCCTTCCGACGGCTAGCCGAGTC
 -5 A H A A L A H L L T A S R P L S G I P E -
 TGGCGATGGTCATGGCCGAAGAGTAGGGAAGAGGCTGGGTTCGAACCACCGCAAAGCTT
 2701 -----+-----+-----+-----+-----+ 2760
 ACCGCTACCGTACCGGCTTCTCATCCCTTCTCCGACCCAAAGCTTGGTGGCGTTTCGAA
 -5-< T A I T M -
 TGATTGCCGCTTTTTTCAGGGGAAGTTGATGCGAAGTCGCCGAGCGGCGGAACGTGCTGAT
 2761 -----+-----+-----+-----+-----+ 2820
 ACTAACGGCGAAAAAGTCCCCTTCAACTACGCTTCAGCGGCTCGCCGCTTGCACGACTA
 GTATGGGGGGCGGGAGGAGCCTGCGGGGTTCTAGGAGCCGGTCCGCGGCCACGGTGGAGGA
 2821 -----+-----+-----+-----+-----+ 2880
 CATACCCCCCGCCCTCCTCGGACGCCCAAGATCCTCGGCCAGCGCCGTGCCACCTCCT
 -4-* * S G T A A V T S S -
 GGTGCCCAGCTGGGAGCGGGGGTCTTTTCGCCGACGCGGTGGGCTCGATGGTGGGGG
 2881 -----+-----+-----+-----+-----+ 2940
 CCACGGGTGACCCCTCGCCCCCAGAAAAGCGGCTGCGCCAACCCGAGCTACCACGCCCC
 -4 T G L Q S R P T K E G V R N P E I T R P -
 GTCGACGGCCTCTCCGGGGGCACCTTGCCGGTAGACGCCTTCGGGGTCGGAGTCCCGGTC
 2941 -----+-----+-----+-----+-----+ 3000
 CAGCTGCCGAGAGGCCCCCGTGAACGGCCATCTGCGGAAGCCCCAGCCTCAGGGCCAG
 -4 D V A E G P A G Q R Y V G E P D S D R D -
 ATGGGGGAGCAGGAAGAAGACCCGGCGCCGGTACAGACCGCTGTCCGGGTCCGCTTCGGC
 3001 -----+-----+-----+-----+-----+ 3060
 TACCCCTCGTCTCTTCTTCTGGGCCGCGCCATGTCTGGCGACAGGCCAGGCGAAGCCG
 -4 H P L L F F V R R R Y L G S D P D A E A -
 GTCGCCCCCGAGTTTCGATGTAGCCGATCATGCGGCCGTGCGGGCGTAGCGCGGCTTGTT
 3061 -----+-----+-----+-----+-----+ 3120
 CAGCCGGGGCTCAAGCTACATCGGCTAGTACGCCGGCAGCGCCCGCATCGCGCCGAACAA
 -4 D A G L E I Y G I M R G D R A Y R P K N -

CTTGCGCCGGGGGTCTTGTCCAGGGCCTGGCGGACGTAGTCGAGTCCCTCGGGATCTTC
 3121 -----+-----+-----+-----+-----+ 3180
 GAACGCGGCCCCCAGAACAGGTCCCGGACCGCCTGCATCAGCTCAGGGAGCCCTAGAAG
 -4 K R R P T K D L A Q R V Y D L G E P D E -
 GAGCCACACGACCTTCGCCTCGTGAACGAGATCGCTGTGGTCACTAGCGAGCTCATGGC
 3181 -----+-----+-----+-----+-----+ 3240
 CTCGGTGTGCTGGAAGCGGAGCACTTGCTCTAGCGACAGCCAGTCATCGCTCGAGTACCG
 -4-< L W V V K A E H V L D S D T L L S S M -
 GGCGACCTCTCCTTCGTCGGCGTGCACCGGTGGGGAAGCGGTGCCTGCGTGATGTGTGT
 3241 -----+-----+-----+-----+-----+ 3300
 CCGCTGGAGAGGAAGCAGCCGCACGTGGCCACCCCTTCGCCACGGACGCACTACACACA
 TCGTCTGCGGCGGTGGGCCGAGTGGTGCGGACCGCCCGTGGTGCCGGTTCTCGGCCAAA
 3301 -----+-----+-----+-----+-----+ 3360
 AGCAGACGCCGCCACCCGGCGTCAACACGCCTGGCGGGCACCACGGCCAAGAGCCGGTTT
 GCACGGGCAGGTACGTCTGGGGCACTCACATCGTAGATGGGGTCCGCTTCCGCAGGGCA
 3361 -----+-----+-----+-----+-----+ 3420
 CGTGCCCGTCCATGCAGGACCCCGTGAGTGTAGCATCTACCCAGGCGAAGGCGTCCCGT
 GTGCCTCCGGTCGGAGGACGTTTCATTGTCGGCTGCCAGAGCGAGGTTGGGGTAGAACTT
 3421 -----+-----+-----+-----+-----+ 3480
 CACGGAGGCCAGCCTCCTGCAAGTAAGCAGCCGACGGTCTCGCTCCAACCCCATCTTGAA
 -3-* * E D A A L A L N P Y F K -
 CCGGCCGTGGATTGATCATGTTCGGCAGGTGAGGCGAGGCCCACTTCTGGCGGACCCG
 3481 -----+-----+-----+-----+-----+ 3540
 GGCCGGCAACCTAAACTAGTACAGCCGTCCACTCCGCTCCGGGTGAAGGACCGCTGGGC
 -3 R G N S K I M D A P S A L G V E Q R V R -
 GGTGGCGAAGGCACGGGCGGTCCCGGGGCGGATGCCTTCACTGTGTGCGCACCAGGTGCT
 3541 -----+-----+-----+-----+-----+ 3600
 CCACCGCTTCCGTGCCCGCCAGGGCCCCGCTACGGAAGTGACACACGCGTGGTCCACGA
 -3 T A F A R A T G P R I G E S H A C W T S -
 GTAGGACGTGTAGAGAAGGCCCTGTTTCGACGCGTAGCTCGCTGTTCTCGGGGTCGTGGAG
 3601 -----+-----+-----+-----+-----+ 3660
 CATCCTGCACATCTCTTCCGGGACAAGCTGCGCATCGAGCGACAAGAGCCCCAGCACCTC
 -3 Y S T Y L L G Q E V R L E S N E P D H L -
 GCAGCACTCGGCGAGGAAGCGGCCGATGTGGTCTCGGTGTTTCGCGTATGCGCTGGTGGC
 3661 -----+-----+-----+-----+-----+ 3720
 CGTCGTGAGCCGCTCCTTCGCCGGCTACACCAGGAGCCACAAGCGCATACGCGACCACCG
 -3 C C E A L F R G I H D E T N A Y A S T A -
 GATGCGGACCCGGTCCGGGGCCGGCGAGTGTGTGCGGGTGGCGAGGTAGCGGCGGGCCCC
 3721 -----+-----+-----+-----+-----+ 3780
 CTACGCCTGGGCCAGCCCCGGCGCTCACACAGCGCCACCGCTCCATCGCCGCCCGGGG
 -3 I R V R D P G A L T D R T A L Y R R A G -
 TTCGGTGAGCCAGTGCAGGATCCCGGGGCCCTCGTCTTGACGAGTTCGACAGCCAGGTT
 3781 -----+-----+-----+-----+-----+ 3840
 AAGCCACTCGGTACGTCCTAGGGCCCCGGGAGCAGGACCTGCTCAAGCTGTGCGTCCAA
 -3 E T L W H L I G P G E D Q V L E V A L N -
 GTCGATCTTGCGTTCGTGCGGGACGATCCGTTTGAAGGGCAGGAGGCGGATGCGGCGCCA
 3841 -----+-----+-----+-----+-----+ 3900
 CAGCTAGAACGCAAGCAGCCCCTGCTAGGCAAGCTTCCCGTCTCCGCCTACGCCGCGGT
 -3 D I K R E D P V I R E F P L L R I R R W -
 GAAGGCGAAGCCCGCGGTGGAGACCTCGGGGCGGTGGTTGCCAGCAGCCACAGCTTGTG
 3901 -----+-----+-----+-----+-----+ 3960
 CTTCCGCTTCGGCGGCCACCTCTGGAGCCCCGCCACCAACGGGTGCTCGGTGTGAACAC

-3 A K E G G A R K W R Y G D W S Y W G L G -
 CTCCACGTGCCGGAAGTGGTCACGGTAGAGACGGACGAAGAGCTTGGCGTTGCCGCGGTC
 4741 -----+-----+-----+-----+-----+-----+-----+ 4800
 GAGGTGCACGGCCTTGACCAGTGCCATCTCTGCCTGCTTCTCGAACCGCAACGGCGCCAG
 -3 E V H R F Q D R Y L R V F L K A N G R D -
 GGTCAGGCTGGCGGGAATCTCGCCCGCCTCCAGGCGGTTCGCGGCGACGGGGGCTCGGG
 4801 -----+-----+-----+-----+-----+-----+-----+ 4860
 CCAGTCCGACCGCCCTTAGAGCGGGCGGAGGTCCGCCAGCGCCGCTGCCCCCGGAGCCC
 -3 T L S A P I E G A E W A T A A V P A E P -
 AGCGGCCTGGACAGGGAGGAGCGGCGCTGGGGCCGGGGTGGTTTCGAGGGCCAGCATCTG
 4861 -----+-----+-----+-----+-----+-----+-----+ 4920
 TCGCCGACCTGTCCCTCCTCGCCGCGACCCCGGCCACCAAAGCTCCCGGTCTAGAC
 -3 A A Q V P L L P A P A P T T E L A L M Q -
 CTGAGCGGCGGCAGTTGCGTCAAAGCGAGGGCCCTCGGCGCTGCTGCTCATGGACGTCTT
 4921 -----+-----+-----+-----+-----+-----+-----+ 4980
 GACTCGCCCGCTCAACGCAGTTTCGCTCCCGGAGCCGCGACGACGAGTACCTGCAGGA
 -3-< Q A A A T A D F R P G E A S S S M -
 TCGAGATGGAGCGGTTCGGGCGGTCCCGCTGCGGGAACGGCATGAATGATCTTCCCGGTG
 4981 -----+-----+-----+-----+-----+-----+-----+ 5040
 AGCTCTACCTCGCCAGCCCGCCAGGGGCGACGCCCTTGCCGTACTTACTAGAAGGGCCAC
 CGGACAGAGTGCCAGGGGCGAGCGCATGTGCGGGGGACAACGGCCCGTTTCGGACGAGGG
 5041 -----+-----+-----+-----+-----+-----+-----+ 5100
 GCCTGTCTACGGTCCCCGTGCGTACACGCCCCCTGTTGCCGGGCAAAGCCTGCTCCC
 CCGGCCGACGGGGGAAGCAGGGGCGGCAACCGGGTGGCGGGGCGGCGTGAGCGAGGGC
 5101 -----+-----+-----+-----+-----+-----+-----+ 5160
 GGCCGGTGCCCCCTTCGTCCCCGCGCTTGGCCACCGCCCCGCGCACTCGCTCCCC
 ACGAGCGGCCCGGTACGGGGGAAGGGCTCGTCTCTCCGTGGGGGCGGCACGTTGTGGTCC
 5161 -----+-----+-----+-----+-----+-----+-----+ 5220
 TGCTCGCCGGGCCATGCCCCCTTCCCGAGCAGAGAGGCACCCCGCGTGCAACACCAGG
 TCGTCCGTGAGCTTGGCTTTCAGCCTCCTGACCCCCAATAAGGCGAAAGCTGCTG
 5221 -----+-----+-----+-----+-----+-----+-----+ 5280
 AGCAGGCAGTCGAACGCAGACCGAAGTCGGAGGACTGGGGGTATTCCGCTTTCGACGAC
 GTCAAGCATCTTTCTGTGACACTCGGCGAGGGACTGAAGGGACTGTCTTTCGGAATGAGTG
 5281 -----+-----+-----+-----+-----+-----+-----+ 5340
 CAGTTCGTAGAAAGCACTGTGAGCCGCTCCCTGACTTCCCTGACAGAAAGCCTTACTCAC
 TAGGGGGTTGTGCGGTGGGGACCGCGCCTCGACTCCCCGCGGACGGGATCTGTTGCGTTC
 5341 -----+-----+-----+-----+-----+-----+-----+ 5400
 ATCCCCCAACAGCCACCCCTGGCGCGGAGCTGAGGGGCGCCTGCCCTAGACAAGCCAG
 GGTCCCTTGGGTCCCTCCCCGGATCGCGGCAGGGACCCAAGGGGGCGGTGCGGCGGGCGG
 5401 -----+-----+-----+-----+-----+-----+-----+ 5460
 CCAGGGAACCCAGGGAGGGGCTAGCGCCGTCCCTGGGTCCCCCGCCACGCCGCCCGCC
 TCGGTGAGGGGCCCCGGTGGAGGGACTGAGGGTCTGTATGGAGCGATAAGAGGGTCTGAA
 5461 -----+-----+-----+-----+-----+-----+-----+ 5520
 AGCCACTCCCCGGGGCCACCTCCCTGACTCCCGACATACCTCGCTATTCTCCAGACTT
 GGGGCGGAGAGAGTTTCGGTCCCTGCGTTGAGTCCCTGGTCATCACCGCAGGTTCAGAGGG
 5521 -----+-----+-----+-----+-----+-----+-----+ 5580
 CCCCCGCTCTCTCAAAGCCAGGGACGCAACTCAGGGACCAGTAGTGGCGTCCAGTCTCCC
 GTTTTGAGGGGTGAAAAAGGGACTGAAGGGACTCAACTTCCCCATTATGAGCTGAGTAGA
 5581 -----+-----+-----+-----+-----+-----+-----+ 5640
 CAAAACTCCCCACTTTTCCCTGACTTCCCTGAGTTGAAGGGTAATACTCGACTCATCT


```

TCAACTTCCGCAGGCTGGCCAGTGGCCGCCGCGAGCCGCGGACGAGCCGCTGAAGCGCC
-2      L K A S D R S P A A L G A L L G D F A A -

CCGTCAACGAGGTCATCACCGCAGCGATGGACGATGTCCTGCGCAGTGGAGCGGACCCCG
7321    -----+-----+-----+-----+ 7380
GGCAGTTGCTCCAGTAGTGGCGTTCGCTACCTGCTACAGGACGCGTCACCTCGCCTGGGGC
-2      V N E V I T A A M D D V L R S G A D P A -

CGAAGGCCTTCGCCCAGCGCGGCGTGGCCGCCAGCAACTGCTCGATGCCTACAACGCCC
7381    -----+-----+-----+-----+ 7440
GCTTCCGGAAGCGGCTTCGGCCGACCGGCGGGTCGTTGACGAGCTACGGATGTTGCGGG
-2      K A F A E A G V A A Q Q L L D A Y N A R -

GGAACCGCTCCGGATCCGGGACCCCTCCGCCGTCTGAGATCCGGTACCGGGGCACAGGG
7441    -----+-----+-----+-----+ 7500
CCTTGGCGAGGCCTAGGCCCTGGGGGAGGCGGCAGACTCTAGGCCATGGCCCCGTGTCCC
-2-*    N R S G S G T P S A V * -

GCGCCGCCCGCCCGCTTTCCCGGCGGGGCACTGGCCGGGGGACATGCTCTCCCGCCCCCGG
7501    -----+-----+-----+-----+ 7560
CGCGGCGGCGGGCGAAAGGGCCGCCCGTGACCGGCCCCCTGTACGAGAGGGCGGGGGCC

CAGGACGTAGGGTCAACCCGCTGCGCCTTCAGGTGGCGGCGCAGATACTCACCGGTCAG
7561    -----+-----+-----+-----+ 7620
GTCCTGCATCCCAGTTGGGCGGACGCGGAAGTCCACCGCCGCGTCTATGAGTGGCCAGTC
-1-*    * G A Q A K L H R R L Y E G T L -

GGAGGAATCCGCGGCGAGCAGGTCTTTCGGTGTGCCGGTGAAGACGATCTCGCCGCCCTC
7621    -----+-----+-----+-----+ 7680
CCTCCTTAGGCGCCGCTCGTCCAGGAAGCCACACGGCCACTTCTGCTAGAGCGGCGGGAG
-1      S S D A A L L D K P T G T F V I E G G E -

CCGTCCCCCGTCGGGACCCAGGTTCGATGATCCAGTTCGGCCTGCTGCACCACATCGAGGTT
7681    -----+-----+-----+-----+ 7740
GGCAGGGGGCAGCCCTGGGTCCAGCTACTAGGTACGCCGACGACGTGGTGTAGCTCCAA
-1      R G G D P G L D I I W D A Q Q V V D L N -

GTGCTCGATGACCACGACGGTGTTCGCCGCCCTCGACGAGCCCGTCCAGGAGCTTCAGCAG
7741    -----+-----+-----+-----+ 7800
CACGAGCTACTGGTGTGCCACAAGGGCCGGAGCTGCTCGGGCAGGTCTCGAAGTCGTC
-1      H E I V V V T N G A E V L G D L L K L L -

GGTGTCAACGTCCGACATGTGCAGCCCGGTGGTGGGCTCGTCCAGGACATAGACCGTGCC
7801    -----+-----+-----+-----+ 7860
CCACAGTTGCAGGCTGTACACGTCCGGGCCACCAACCGAGCAGGTCTGTATCTGGCACGG
-1      T D V D S M H L G T T P E D L V Y V T G -

CGTGCGGTGCAGCTGGTTCGGCAAGTTTGATCCGCTGCAGTTACCGCCGGAGAGGCTGGA
7861    -----+-----+-----+-----+ 7920
GCACGCCACGTTCGACCAGCCGTTCAAAGTAGGCGACGTCAAGTGGCGGCCCTCTCCGACCT
-1      T R H L Q D A L K I R Q L E G G S L S S -

AAGCGGCTGGCCCAGGCTGAGGTACCCAAGACCGACGTTCGACGAGAGCGCGCAGTTTCGG
7921    -----+-----+-----+-----+ 7980
TTCGCCGACCGGGTCCGACTCCATGGGTCTGCGTGCAGCTGCTCTCGCGCGTCAAAGCC
-1      L P Q G L S L Y G L G V D V L A R L K P -

CAGCAGGGCCTTCTCGGTGAAGAACTCGACGGCCTCGTCGGCGGGCAGCTCCAGGACGTC
7981    -----+-----+-----+-----+ 8040
GTCGTCCCGGAAGAGCCACTTCTTGAGCTGCCGGAGCAGCCGCCCGTCGAGGTCTGTCAG
-1      L L A K E T F F E V A E D A P L E L V D -

CGCGATCGACTTCCCGCGAAGCTGGTGTCTCCAGGACCTCGGGCTTGAAGCGGCGCCCCCTC
8041    -----+-----+-----+-----+ 8100
GCGCTAGCTGAAGGGCGCTTCGACCACGAGGTCTTGAGAGCCCGAAGTTCGCCGCGGGGAG
-1      A I S K G R L Q H E L V E P K F R R G E -

```

```

ACAGACACCGCAGTGGTGGTCACCGGATCCATGAAGGCCAGCTCGGTGATGATGACCCC
8101 -----+-----+-----+-----+-----+-----+ 8160
TGTCTGTGGCGTCACGCACCACTGGCCTAGGTACTTCCGGTCGAGCCACTACTACTGGGG
-1 C V G C H T T V P D M F A L E T I I V G -

GCGGCCCTGGCACTCCTCGCACGACCCCTTGGAGTTGAAGCTGAACAGCGAGGCGTTCGC
8161 -----+-----+-----+-----+-----+-----+ 8220
CGCCGGGACCGTGAGGAGCGTGTCTGGGGAACCTCAACTTCGACTTGTCTCGCTCCGCAAGCG
-1 R G Q C E E C S G K S N F S F L S A N A -

GCCGGTCTCCTTCGCGAACAGCTTGCAGCAGCGGGTCCATCAGGCCGAGGTAGGAGACCGG
8221 -----+-----+-----+-----+-----+-----+ 8280
CGGCCAGAGGAAGCGCTTGTCTGAACGCGTCTGCCAGGTAGTCCGGCTCCATCCTCTGGCC
-1 G T E K A F L K R L P D M L G L Y S V P -

TGTGGAGCGCGACGAGGCGGCGATCGCGGACTGGTCTGACAAAGACCGCGTCCGGGTGCGC
8281 -----+-----+-----+-----+-----+-----+ 8340
ACACCTCGCGCTGCTCCGCCGCTAGCGCCTGACCAGCTGTTTCTGGCGCAGCCCCACGCG
-1 T S R S S A A I A S Q D V F V A D P H A -

CTCCATGAATGCCCCGGAGATCAGGCTGCTCTTGCCGGAACCCGCCACCCCGGTACCCGC
8341 -----+-----+-----+-----+-----+-----+ 8400
GAGGTACTTACGGGGCCTCTAGTCCGACGAGAACGGCCTTGGGCGGTGGGGCCAGTGGCG
-1 E M F A G S I L S S K G S G A V G T V A -

GGTCAGCACACCGGTGGGCACGGCCACGGAGACCTGCTTCAGGTTGTGGAGATCCGCGTT
8401 -----+-----+-----+-----+-----+-----+ 8460
CCAGTCGTGTGGCCACCCGTGCCGGTGCCTCTGGACGAAGTCCAACACCTCTAGGCGCAA
-1 T L V G T P V A V S V Q K L N H L D A N -

CTCCACGGTCAGCTCCCCCGTGGGCGGGCGGACCTCCTCCTTACGCGGGGCCCCCGCGC
8461 -----+-----+-----+-----+-----+-----+ 8520
GAGGTGCCAGTCGAGGGGGCACCCGCCCGCCTGGAGGAGGAAGTGCGCCCGGGGGCGGC
-1 E V T L E G T P P R V E E K V R A G R R -

CAGAGCCTCCCCGGTCCGGGTCTTCGCCTTCCGCAGCTTCGCGAAGGACCCCTCGAACAC
8521 -----+-----+-----+-----+-----+-----+ 8580
GTCTCGGAGGGGCCAGGCCCAGAAGCGGAAGGCGTCAAGCGCTTCTTGGGGAGCTTGTG
-1 L A E G T R T K A K R L K A F S G E F V -

GATCTCGCCCCCGTGCACTCCCGCCCCGGGACCGACATCGACGATGTGGTTCGGCGATCTC
8581 -----+-----+-----+-----+-----+-----+ 8640
CTAGAGCGGGGGCACGTGAGGGCGGGGCCCTGGCTGTAGCTGCTACACCAGCCGCTAGAG
-1 I E G G H V G A G P G V D V I H D A I E -

GATCACATcGGGGTTCGTGCTCGACGACCAGCACGGTGTTCCTTGTTCGCGCAGCGCGCG
8641 -----+-----+-----+-----+-----+-----+ 8700
CTAGTGTAgCCCCAGCACGAGCTGCTGGTCTGCCACAAGGGGAACAGCGCTCGCGCGC
-1 I V D P D H E V V L V T N G K D R L A R -

CAGCAGGTCGTTGAGCCGCCCCACGTGCGCGGGGTGCAGGCCGATGCTGGGCTCGTCGAA
8701 -----+-----+-----+-----+-----+-----+ 8760
GTCGTCCAGCAACTCGGCGGGGTGCAGCGCGCCACGTCCGGCTACGACCCGAGCAGCTT
-1 L L D N L R G V D R P H L G I S P E D F -

GATGTACGTGAGCCCGGCCAGACCACTGCCGAGGTGGCGCACCATCTTTCAGCCGCTGCCC
8761 -----+-----+-----+-----+-----+-----+ 8820
CTACATGCACTCGGGCCGGTCTGGTGACGGCTCCACCGCGTGGTAGAAGTCGGCGACGGG
-1 I Y T L G A L G S G L H R V M K L R Q G -

CTCGCCCCCGAGAGGTTCGGCCGTGGGCCTGTCCAGGGTCAGGTAGCCGAGCCCGATGGA
8821 -----+-----+-----+-----+-----+-----+ 8880
GAGCGGGGGGCTCTCCAGCCGGCACCCGGACAGGTCCAGTCCATCGGCTCGGGCTACCT
-1 E G G S L D A T P R D L T L Y G L G I S -

```

8881 CACGATCCGCTCCAGGGCCGTGCGCGCGGCTTTTCGCGAGAGGGGCAGCGCCGGCTCCGT 8940
 -----+-----+-----+-----+-----+-----+-----+
 -1 GTGCTAGGCGAGGTCCCGGCACGCGCGCCGAAAGCGCTCTCCCCGTCGCCGGCCGAGGCA
 V I R E L A T R A A K A L P A A A P E T -
 8941 GACGCCGGCGAGCACCTCCGTGAGGTGCGCGACCTCCATGCTCGAGTAGTCGGCGATGTT 9000
 -----+-----+-----+-----+-----+-----+-----+
 -1 CTGCGGCCGCTCGTGGAGGCACTCCAGCGCCTGGAGGTACGAGCTCATCAGCCGCTACAA
 V G A L V E T L D R V E M S S Y D A I N -
 9001 CTTGCCGTCGATCCGGACGTGAGCGCGCGCGGCGTTGAGCCGCGCGCCCCGGCAGGAGGG 9060
 -----+-----+-----+-----+-----+-----+-----+
 -1 GAACGGCAGCTAGGCCTGCAGCTCGCGCCGCCGCAACTCGGCGCGCGGGGCCGTCCTCCC
 K G D I R V D L A A A N L R A G R C S P -
 9061 ACAGACTCCGTCgGTGACGAAACGTTTCGATGACCTCGCGCTTTCGGTTCGcTCAGCGCGCT 9120
 -----+-----+-----+-----+-----+-----+-----+
 -1 TGTCTGAGGCAGcCACTGCTTTGCAAGCTACTGGAGCGCGAACGCCAGCgAGTCGCGCGA
 C V G D T V F R E I V E R K R D S L A S -
 9121 GAGGTGCGCGCTTTCGAGTTGAgCCGCTCGAACCGGTTCGGCcAACCCCTCGTAGTTTCGTCTG 9180
 -----+-----+-----+-----+-----+-----+-----+
 -1 CTCCAGCGCGAACTCCAACtcGGCGAGCTTGGCCAGCCGgTTGGGGAGCATCAAGCAGAC
 L D R K L N L R E F R D A L G E Y N T Q -
 9181 GAACTCGGTGCTCTTGGTCTTCAGCGTcACCTTCCCGCCGGTGcCGCGCAGCAGCGTGTC 9240
 -----+-----+-----+-----+-----+-----+-----+
 -1 CTTGAGCCACGAGAACCAGAAGTCGCAgTGAAGGGCGGCCACgGCGCGTCGTTCGCACAG
 F E T S K T K L T V K G G T G R L L T D -
 9241 CAGCTCCTCGGCGCTGTACTCGGCGATCGGCTTGGCCGGATCCAGACGGCCGGACTTCGC 9300
 -----+-----+-----+-----+-----+-----+-----+
 -1 GTCGAGGAGCCGCGACATGAGCCGCTAGCCGAACCGGCCTAGGTCTGCCGGCCTGAAGCG
 L E E A S Y E A I P K A P D L R G S K A -
 9301 CCAGATCTGCCAGTCCGGGCTACCCACCTTGTACTCGGGGAAAAGGACCGCCCCGTCGTC 9360
 -----+-----+-----+-----+-----+-----+-----+
 -1 GGTCTAGACGGTCAGGCCCGATGGGTGGAACATGAGCCCCTTTTCTGGCGGGGCAGCAG
 W I Q W D P S G V K Y E P F L V A G D D -
 9361 CAGGGACTTCGAGCGGTCCAGCATCTTGTCCAGGTGAGGGCGATGCTCTGGCCGAGACC 9420
 -----+-----+-----+-----+-----+-----+-----+
 -1 GTCCCTGAAGCTCGCCAGGTCTAGAACAGGTCCAGCTCCCGCTACGAGACCGGCTCTGG
 L S K S R D L M K D L D L A I S Q G L G -
 9421 GTCGCAGTCCGGGCACATGCCCTGGGGGTCGTTGAACGAGAACCGGAGACGCCGAGCGA 9480
 -----+-----+-----+-----+-----+-----+-----+
 -1 CAGCGTCAGGCCCGTGTACGGGACCCCCAGCAACTTGCTCTTTCGCGCTCTGCGGCTCGCT
 D C D P C M G Q P D N F S F A S V G L S -
 9481 GGACGGCCCCGTCGTCCTTCGTCGTGCCGAACCGTGCGAACAGGGCCCCGGATCATCGGCTG 9540
 -----+-----+-----+-----+-----+-----+-----+
 -1 CCTGCCGGGCAGCAGGAAGCAGCACGGCTTGGCACGCTTGTCCCGGCCCTAGTAGCCGAC
 S P G D D K T T G F R A F L A R I M P Q -
 9541 TACGTCCGTCATGGTCCCCACCGTGGACCGGGCGTTGCCCCCACGGGCTTCTGGTCGAC 9600
 -----+-----+-----+-----+-----+-----+-----+
 -1 ATGCAGGCAGTACCAGGGGTGGCACCTGGCCCCGCAACGGGGGGTGCCCGAAGACCAGCTG
 V D T M T G V T S R A N G G V P K Q D V -
 9601 GATCACCGGGGTGGTGAAGTTCTCGATCGCCTCGGCCTGAGGACGTTCTGACTTCGGAAG 9660
 -----+-----+-----+-----+-----+-----+-----+
 -1 CTAGTGGCCCCACCACTCCAAGAGCTAGCGGAGCCGACTCCTGCAAGCATGAAGCCTTC
 I V P T T L N E I A E A Q P R E Y K P L -
 CTGGTTGCGGATGTACCAGCTGAAGGTGGAGTTCAGCTGTCGCTGGGCCTCCACGGCCAC

```

9661 -----+-----+-----+-----+-----+-----+ 9720
GACCAACGCCTACATGGTTCGACTTCCACCTCAAGTCGACAGCGACCCGGAGGTGCCGGTG
-1      Q N R I Y W S F T S N L Q R Q A E V A V -

CGTGTCTGAAGACGATCGACGACTTGCCCGAACCCGAGACCCCGTGAAGACCGTGATCTG
9721 -----+-----+-----+-----+-----+-----+ 9780
GCACAGCTTCTGCTAGCTGCTGAACGGGCTTGGGCTCTGGGGGCACTTCTGGCACTAGAC
-1      T D F V I S S K G S G S V G T F V T I Q -

GTTGCGGGGAATCGTCAGGGAGACATCTTTGAGGTTGTGGATCCGCGCGCCCGCGATGCG
9781 -----+-----+-----+-----+-----+-----+ 9840
CAACGCCCCCTTAGCAGTCCCTCTGTAGAACTCCAACACCTAGGCGCGCGGGGCGCTACGC
-1      N R P I T L S V D K L N H I R A G A I R -

GATGCCGTCTCCCGGGCCGGATGTTTTTCCCGCGCCGGCGGTGGGGTCCGGTGACGCTCAC
9841 -----+-----+-----+-----+-----+-----+ 9900
CTACGGCAGAGGGCCCGGCTACAAAAAGGGCGCGGCCGCCACCCAGCCACTGCGAGTG
-1-<    I G D G P G S T K G A G A T P D T V S M-

AGAGTTTCTCCTGGCTTCCGTACATGATTTACCGTGTCTAGCCGGGCAAACCGGCGGAA
9901 -----+-----+-----+-----+-----+-----+ 9960
TCTCAAAAGGAGGACCGAAGGCATGTACTAAATGGCACAGTCGGCCCGTTTGGCCGCCTT

CGGTAACCACCTAGCTTGTACTCAGGAGGTGTCCGGGGTCTTCTCCTCCCGTGCTGACTT
9961 -----+-----+-----+-----+-----+-----+ 10020
GCCATTGGTGGATCGAACATGAGTCCTCCACAGGCCCCAGAAGAGGAGGGGCACGACTGAA
0-*      * S T D P T K E E R A S K -

GGGGGCCGGCCCCGCGGACAGGGCCGGCTCCGTGTTCCACCCCGCCAGCCGATCCCCCGG
10021 -----+-----+-----+-----+-----+-----+ 10080
CCCCCGGCCGGGCGGCTGTCCCGGCCGAGGCACAAGGTGGGGCGGTCCGGCTAGGGGGGC
0      P A P G G S L A P E T N W G A L R D G R -

CTCCGTCTCGTCTCCTCGAGAACGATCCGGCTGCTCGCCCAGCGCAGGATCGGCGGCGC
10081 -----+-----+-----+-----+-----+-----+ 10140
GAGGCAGAGCAGGAGGAGCTCTTGCTAGGCCGACGAGCGGGTCCGCTCCTAGCCGCCGCG
0      E T E D E E L V I R S S A W R L I P P A -

CGTCACCGAGGTGATGAGGGCGACCAGCACGATGATCGTGAAGGTCACGGTGTCCAGTAC
10141 -----+-----+-----+-----+-----+-----+ 10200
GCAGTGGCTCCACTACTCCCGCTGGTCTGCTACTAGCACTTCCAGTGCCACAGGTCATG
0      T V S T I L A V L V I I T F T V T D L V -

GCCGATACGCAGGCCGACCAGGGCGATCACCACCTCGATCATTCCACGCGAGTTCATCCC
10201 -----+-----+-----+-----+-----+-----+ 10260
CGGCTATGCGTCCGGCTGGTCCCGCTAGTGGTGGAGCTAGTAAGGTGCGCTCAAGTAGGG
0      G I R L G V L A I V V E I M G R S N M G -

CGCTCCGAGCGCCAGCCCCTCGTAGCGGCTCATCCCGCCACTACGGGCGGCGACGTACGC
10261 -----+-----+-----+-----+-----+-----+ 10320
GCGAGGCTCGCGGTCCGGGAGCATCGCCGAGTAGGGCGGTGATGCCCCGCGCTGCATGCG
0      A G L A L G E Y R S M G G S R A A V Y A -

ACCGGCGAACTTGCCGAAAGTGGCCACCAACAGCACCCCCGAGGCCCGTGAGCAGCACCGA
10321 -----+-----+-----+-----+-----+-----+ 10380
TGGCCGCTTGAACGGCTTTTACCAGGTGGTTGTGCTGGGGCTCCGGGCACTCGTCTGGCT
0      G A F K G F T A V L L V G L G T L L V S -

CGGCTCCGCGAGTGCGGTCCAGGTCCATGCGAAGCCCCCACTGCCAGGAACACCGGTGC
10381 -----+-----+-----+-----+-----+-----+ 10440
GCCGAGGCGCTCACGCCAGTCCAGGTACGCTTCCGGGTGTGACGGGTCCTTGTGGCCACG
0      P E A L A T L D M R L G V S G L F V P A -

GAACACGGCCATGACCAGCGTGCGCAGCGGGGCGAGCCGTACCGGGGCGATGTGCCTCAG
10441 -----+-----+-----+-----+-----+-----+ 10500
CTTGTGCCGGTACTGGTTCGACGCGTCGCCCCGCTCGGCATGGCCCCGCTACACGGAGTC

```

0 F V A M V L T R L P A L R V P A I H R L -
 CAGGGTCGCACCGGCCACGAACGCCCCGAACAACGCCTCCATCCCGGCCCGCGCGGTACG
 10501 -----+-----+-----+-----+-----+ 10560
 GTCCACAGCGTGGCCGGTGCTTGCGGGCTTGTTGCGGAGGTAGGGCCGGCGGCCAGTC
 0 L T A G A V F A G F L A E M G A A A T L -
 CGCCCCGTACAGGACGACCACGGCCACGCCGACGGTGACGGCCGATACGGGGACCCGGCT
 10561 -----+-----+-----+-----+-----+ 10620
 GCGGGGCATGTCTGCTGGTGCCGGTGCGGCTGCCACTGCCGGCTATGCCCTGGGCGGA
 0 A G Y L V V V A V G V T V A S V P V R S -
 GTCACCCGTACGGGACAGCCGCTGCCGATCGGGCCGCCACCGCACACGCCGCGGCGAC
 10621 -----+-----+-----+-----+-----+ 10680
 CAGTGGGCATGCCCTGTCCGCGGACGGCTAGCCCGCGGGTGGCGTGTGCGGCGCCGCTG
 0 D G T R S L R R G I P G G V A C A A A V -
 GAAGACGGTTCGTCCAGGCCATCGTGGTCAGGACCACGGGGCCCCCGGCCGCCCCACTCGC
 10681 -----+-----+-----+-----+-----+ 10740
 CTTCTGCCAGCAGGTCCGGTAGCACCAGTCTTGGTGCCCGGGGGCGGCGGGGTGAGCG
 0 F V T T W A M T T L V V P G G A A G S A -
 CAGCGCCGTCACCAGAGCGAGCAGCAGCCAGCCACCGCGTCGTCGAACACCGCTGCCGC
 10741 -----+-----+-----+-----+-----+ 10800
 GTCGCGGCAGTGGTCTCGCTCGTCTCGGTGCGGTGGCGCAGCAGCTTGTGGCGACGGCG
 0 L A T V L A L L L W G V A D D F V A A A -
 GATGAGCAGCTGGCCGACGTTGCGGTGCGTCAGATTAGGTTCGGCGAGCGTCTTGGCGAT
 10801 -----+-----+-----+-----+-----+ 10860
 CTACTCGTCGACCGGCTGCAACGCCACGCAGTCTAAGTCCAGCCGCTCGCAGAACCGCTA
 0 I L L Q G V N R H T L N L D A L T K A I -
 CACCGGGAGGGCCGTGACACACATCGCGACCCCGAGGAACAGCGCGAAGACGCCCCGCTC
 10861 -----+-----+-----+-----+-----+ 10920
 GTGGCCCTCCCGGCACTGTGTGTAGCGCTGGGGCTCCTTGTGCGGCTTCTGCGGGCGAG
 0 V P L A T V C M A V G L F L A F V G R E -
 TCCGGAGTCCGCGAGCAGCGAGGCGGGCACCAGGTAGCCGGTGGCGATGCCAGCCCCAG
 10921 -----+-----+-----+-----+-----+ 10980
 AGGCCTCAGGCGCTCGTCTCGCTCCGCCCCGTGGTCCATCGGCCACCGCTACGGGTGCGGGTC
 0 G S D A L L S A P V L Y G T A I G L G L -
 AGGAATCAGAAGACCCGCCAGGCTGACCCGGGCGGCCAGACCCCGCGCTTGCAGGAT
 10981 -----+-----+-----+-----+-----+ 11040
 TCCTTAGTCTTCTGGGCGGTCCGACTGGGCCCCGCGGTCTGGGGGCGGAACGCGTCCTA
 0 P I L L G A L S V R A A L G G R K R L I -
 CCGGGGTTCGAAC TGGGCACCTGCGATGGCCACCAGCAGAAGGACGCCGAAC TGGCAGAA
 11041 -----+-----+-----+-----+-----+ 11100
 GGCCCCCAGCTTGACCCGTGGACGCTACCGGTGGTCTCTTCTGCGGCTTGACCGTCTT
 0 R P D F Q A G A I A V L L L V G F Q C F -
 CGCGTCGAGCAGGTGCGCCTGCGAGATGTCTCGGAAACAGCCTGCCGAAAGTCCCGG
 11101 -----+-----+-----+-----+-----+ 11160
 GCGCAGCTCGTCCACGCGGACGCTCTACAGGAGCCCTTTGTGCGACGGCCCTTTCAGGGCC
 0 A D L L H A Q S I D E P F L R G S L G P -
 CGAGATCTGCCCCAGCAGGGTCGGCCCCGAGCAGTACCCCCGCGGTGAGCTCCCCACCAG
 11161 -----+-----+-----+-----+-----+ 11220
 GCTCTAGACGGGGTCGTCCCAGCCGGGCTCGTCATGGGGGCGCCAGTCGAGGGGTGGTC
 0 S I Q G L L T P G L L V G A T L E G V L -
 CGGCGGCAGACCGATCCGGGTCCCCAGCCGTCCAGACCGTAGGCACAGGCGAGCAGGAG
 11221 -----+-----+-----+-----+-----+ 11280
 GCCGCCGTCTGGCTAGGCCAGGGGTGCGCAGGGTCTGGCATCCGTGTCCGCTCGTCTCTC
 0 P P L G I R T G L R G L G Y A C A L L L -

0 11281 11340
GCCGACCTGGAGCAGGAAGACCGTCCAGCGGCTCCCCCGCCAGCGACGTGGCTGCAG
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
CGGCTGGACCTCGTCCTTCTGGCAGTCGCCGAGGGGCGGGTCGCCGCTGCACCGACGCTC
G V Q L L F V T L P E G G L P S T A A L -
CACAGCCACGTCAGGACCGCGCACCGGGAACCCAGCCCAGCCCCGTCCGTCGACCGGCCA
11341 11400
-----+-----+-----+-----+-----+-----+-----+-----+-----+
GTGTCGGTGCGAGTCTTGGCGCGTGGCCCTTGGGTGCGGTGCGGCAGGCAGCTGCGCCGGT
V A V -
11-* * S R A G P V W G L G D T S A A -
GACCCCCCTGCCTCACCGGTCGCTCGGCCCCCGCCTCATCCCCAGAAGAGCCCGTGCCT
11401 11460
-----+-----+-----+-----+-----+-----+-----+-----+-----+
CTGGGGGGACGGAGTGGCCAGCGAGCCGGGGGCGGAGTAGGGGGTCTTCTCGGGCAGCGA
L G G Q R V P R E A G A E D G L L A R A -
GCAGTGCGGCGCTCTGCTCCATGAGGCGGCCACCACCTTTCCCGGCACGGCGCCGTGCG
11461 11520
-----+-----+-----+-----+-----+-----+-----+-----+-----+
CGTCACGCCGCGAGACGAGGTACTCCGCCGGGTGGTGGAAAGGGCCGTGCCGCGGCACGC
Q L A A S Q E M L R G V V K G P V A G H -
GCCCCGTCGGCGTCGCCCCGAGCGGTGTGCGTTCATGCCGGCCATCTCGTCGGACGCCTCGG
11521 11580
-----+-----+-----+-----+-----+-----+-----+-----+-----+
CGGGCAGCCGCGAGCGGGCGTCGCCACACGCGAGTACGGCCGGTAGAGCAGCCTGCGGAGCC
P G D A D G A A T H T M G A M E D S A E -
AGAACCCTGCCTGGCCCCGGGCCGTGTGCGCGAACTCGTCGGAGGAGACCCCGCCGATCA
11581 11640
-----+-----+-----+-----+-----+-----+-----+-----+-----+
TCTTGGCGACGGACCGGGCCCGGCACAGCCGCTTAGCAGCCTCCTCTGGGGCGGCTAGT
S F R Q R A R A T D A F E D S S V G G I -
GTTTCGACGAAGGACTGCAGGTTCGGAGTCCGCGGTGTTGGAGATCTTCCGGGCCTGCCAGA
11641 11700
-----+-----+-----+-----+-----+-----+-----+-----+-----+
CAAGCTGCTTCTGACGTCCAGCCTCAGGCGCCACAACCTCTAGAAGGCCCGACGGTCT
L E V F S Q L D S D A T N S I K R A Q W -
AATAGGAGTCTCTCCGAATGGTGCATGTTCGTAGAACCCGACCAGGAACCTCTAGAACCGGC
11701 11760
-----+-----+-----+-----+-----+-----+-----+-----+-----+
TTATCCTCAGGAGGCTTACCACGTACAGCATCTTCGGCTGGTCTTTCGATCTTCGCCG
F Y S D E S H H M D Y F G V L F E Y F R -
CGTACTCCAGCCGGTAGCGGGCCTCGAACTCCTCGAACGCGCTGGTCTCGTCGACCGACC
11761 11820
-----+-----+-----+-----+-----+-----+-----+-----+-----+
GCATGAGGTTCGGCCATCGCCCGGAGCTTGGAGAGCTTGCAGGACAGAGCAGCTGGCTGG
G Y E L R Y R A E F E E F A S T E D V S -
CGTCCAGGCAGGAGTTGAGCGAGCGCGCTGCCAGCAGTCCGCTGTAGGTGGCGAGGTGCA
11821 11880
-----+-----+-----+-----+-----+-----+-----+-----+-----+
GCAGGTCCGTCCTCAACTCGCTCGCGCGACGGTTCGTAGGCGACATCCACCGCTCCACGT
G D L C S N L S R A A L L G S Y T A L H -
CCCCGGAGGAGAACACCGGGTCGACGAAGCACGCGGCATCCCCGACCAGGGCCATGCCCCG
11881 11940
-----+-----+-----+-----+-----+-----+-----+-----+-----+
GGGGCCTCCTCTTGTGGCCAGCTGCTTCGTGCGCCGTAGGGGCTGGTCCCGGTACGGGC
V G S S S F V P D V F C A A D G V L A M G -
GCGCCCAGAACTTCGTGTTGCTGTACGACCAGTCCTTTCGGACCCGAGCTCGCCGTAGG
11941 12000
-----+-----+-----+-----+-----+-----+-----+-----+-----+
CGCGGGTCTTGAAGCACAACGACATGCTGGTCAGGAACGCCTGGGCCTCGAGCGGCATCC
P A W F K T N S Y S W D K R V R L E G Y -
GGCCCTCGGTCAACCGGGTGGCCTCGGAGAGCTTCTCCGCGATCAGCGGGCAGGCCGCGA
12001 12060
-----+-----+-----+-----+-----+-----+-----+-----+-----+
CCGGGAGCCAGTGGGCCCCAGGAGCCTCTCGAAGAGGCGCTAGTCGCCCGTCCGGCGCT
P G E T V R T A E S L K E A I L P C A A -

[illegible]

12841 -----+-----+-----+-----+-----+-----+-----+ 12900
GAGGAAGAGCGTGTAGCCCGCAGAGTATAAGGGTCTTAGGAGACCGGGCGGGTCCACGA
GCCGCATCTTCGGTATTGCGAAGTCGTGGGCATTCTGCGAGAAGCATGAACCGCGTGGCC
12901 -----+-----+-----+-----+-----+-----+ 12960
CGGCGTAGAAGCCATAACGCTTCAGCACCCGTAAGACGCTCTTCGTACTTGGCGCACCGG
CGGTCTACAGTGGCGTGGAATTTCACTGATTGCGCTGAAGGGCGGCACACGATGAAGGCA
12961 -----+-----+-----+-----+-----+-----+ 13020
GCCAGATGTCACCGCACCTTAAAGTCACTAACGCGACTTCCCGCCGTGTGCTACTTCCGT
10-> M K A -
CTTGTAAGTGTGCGGTGGTTCGGGGACCCGCTGCGCCCGATCAGTTACGCCATGCCGAAG
13021 -----+-----+-----+-----+-----+-----+ 13080
GAACATGACAGCCCAAGCCCTGGGCGGACGCGGGCTAGTCAATGCGGTACGGCTTC
10 L V L S G G S G T R L R P I S Y A M P K -
CAGCTCGTTCGGATCGCCGGAAGCCAGTCTTGAATATGTTCTGGATAATATCCGGAAC
13081 -----+-----+-----+-----+-----+-----+ 13140
GTCGAGCAAGGCTAGCGGCCCTTCGGTCAGGAACCTTATACAAGACCTATTATAGGCCTTG
10 Q L V P I A G K P V L E Y V L D N I R N -
CTCGATATCAAAGAGGTCGCCATTGTCTGTCGGTACTGGGCTCAGGAAATTATTGAGGCA
13141 -----+-----+-----+-----+-----+-----+ 13200
GAGCTATAGTTTCTCCAGCGGTAACAGCAGCCACTGACCCGAGTCTTTAATAACTCCGT
10 L D I K E V A I V V G D W A Q E I I E A -
ATGGGTGACGGCAGCCGTTTCGGTCTGCGCCTCACCTACATACGCCAGGAGCAACCTCTG
13201 -----+-----+-----+-----+-----+-----+ 13260
TACCCACTGCCGTCGGCAAAGCCAGACGCGGAGTGGATGTATGCGGTCCTCGTTGGAGAC
10 M G D G S R F G L R L T Y I R Q E Q P L -
GGCATCGCGCACTGCGTGAAACTGGCCCGAGACTTCCTCGACGAGGACGACTTCGTCCCTC
13261 -----+-----+-----+-----+-----+-----+ 13320
CCGTAGCGCGTGACGCACTTTGACCGGGCTCTGAAGGAGCTGCTCCTGCTGAAGCAGGAG
10 G I A H C V K L A R D F L D E D D F V L -
TACCTAGGCGACATCATGCTGGACGGAGACCTGTCCGCGCAGGCGGGGCACTTCCTCCAC
13321 -----+-----+-----+-----+-----+-----+ 13380
ATGGATCCGCTGTAGTACGACCTGCCTCTGGACAGGCGCGTCCGCCCCGTGAAGGAGGTG
10 Y L G D I M L D G D L S A Q A G H F L H -
ACCCGCCCCCGCGCGGATCGTCTGCGCCAGGTGCCCCGACCCCGGGGCTTCGGGGTG
13381 -----+-----+-----+-----+-----+-----+ 13440
TGGGCGGGGCGGCGGCCTAGCAGCACGCGGTCCACGGGCTGGGGGCCCCGAAGCCCCAC
10 T R P A A R I V V R Q V P D P R A F G V -
ATCGAGCTGGACGGCGAAGGGCGTGTGCTGCGCCTGGTTCGAGAAACCCCGTGAACCGCGC
13441 -----+-----+-----+-----+-----+-----+ 13500
TAGCTCGACCTGCCGCTTCCCGCACACGACGCGGACCAGCTCTTTGGGGCACTTGGCGCG
10 I E L D G E G R V L R L V E K P R E P R -
AGCGACCTCGCGGCGGTGCGCGTGTACTTCTTACCGCGGACGTGCACCGCGCCGTCGAC
13501 -----+-----+-----+-----+-----+-----+ 13560
TCGCTGGAGCGCCCGCAGCCGACATGAAGAAGTGGCGCCTGCACGTGGCGCGGCAGCTG
10 S D L A A V G V Y F F T A D V H R A V D -
GCGATTAGCCCCGAGCCGACGGGGCGAGCTGGAAATCACCGACGCCATCCAGTGGCTGCTG
13561 -----+-----+-----+-----+-----+-----+ 13620
CGCTAATCGGGCTCGGCTGCCCCGCTCGACCTTTAGTGGCTGCGGTAGGTACCGACGAC
10 A I S P S R R G E L E I T D A I Q W L L -
GAGCAGGGCCTGCCGGTTCGAGGCCGGCCGCTACACGGAAGGACACCGGCCCGG
13621 -----+-----+-----+-----+-----+-----+ 13680
CTCGTCCCGGACGGCCAGCTCCGGCCGGCGATGTGCCTGATGACCTTCCTGTGGCCGGCC
10 E Q G L P V E A G R Y T D Y W K D T G R -

```

GTCGAGGACGTCGTGGAGTGCAACCGGCGGATGCTCGGCCGTCTGGCGCTCCAGGTGTCG
13681 -----+-----+-----+-----+-----+-----+ 13740
CAGCTCCTGCAGCACCTCACGTTGGCCGCCTACGAGCCGGCAGACCGCGAGGTCCACAGC
10 V E D V V E C N R R M L G R L A L Q V S -

GGCGAGGTGGACCCGGAGAGCGAACTGGTGGGTGCGGTGGTTCGTTCGAGGAGGGCGCCCGG
13741 -----+-----+-----+-----+-----+-----+ 13800
CCGCTCCACCTGGGCCTCTCGCTTGACCACCCACGCCACCAGCAGCTCCTCCC CGGGGCC
10 G E V D P E S E L V G A V V V E E G A R -

GTGACGCGTTTCGCGGGTCGTGGGACCAGCGGTGATCGGCGCGGGCACGGTCGTTCGAGGAC
13801 -----+-----+-----+-----+-----+-----+ 13860
CACTGCGCAAGCGCCAGCACCTGGTTCGCCACTAGCCGCGCCCGTGCCAGCAGCTCCTG
10 V T R S R V V G P A V I G A G T V V E D -

AGCCAGATCGGACCGTACGCCTCCATCGGCCGGCGCTGCACCGTGCGGGCGTCCC GGCTC
13861 -----+-----+-----+-----+-----+-----+ 13920
TCGGTCTAGCTGGCATGCGGAGGTAGCCGGCCGCGACGTGGCACGCCCGCAGGGCCGAG
10 S Q I G P Y A S I G R R C T V R A S R L -

TCCGACTCCATCGTCCTTGACGACGCCTCGATCCTCGCGGTGAGCGGACTGCACGGCTCG
13921 -----+-----+-----+-----+-----+-----+ 13980
AGGCTGAGGTAGCAGGAAGTCTGCGGAGCTAGGAGCGCCACTCGCCTGACGTGCCGAGC
10 S D S I V L D D A S I L A V S G L H G S -

CTGATCGGAAGGGGCGCGCGGATCGCGCCCGGGGCCCGGGCGAGGCCCGGCACCGGCTG
13981 -----+-----+-----+-----+-----+-----+ 14040
GACTAGCCTTCCCCGCGCGCCTAGCGCGGGCCCGGGCCCGCTCCGGGCGGTGGCCGAC
10 L I G R G A R I A P G A R G E A R H R L -

GTCGTGCGGCACACGTGCAGATCGAGATCGCGGCCTGACGCACCCACCGGAGCACCGGG
14041 -----+-----+-----+-----+-----+-----+ 14100
CAGCAGCCGCTGGTGCACGTCTAGCTCTAGCGCCGACTGCGTGGGTGGCCTCGTGGCCC
10-* V V G D H V Q I E I A A * -

GGGAGGCTCGGCAGGGGCGTCAGGCCGTAAGAAGGGCTGCCGGGGCGGGACGGACCCGCC
14101 -----+-----+-----+-----+-----+-----+ 14160
CCCTCCGAGCCGTCCCCGCGAGTCCGGCATTCTTCCCGACGGCCCCGCCCTGCCTGGGCGG

CCGGCAGCCACAGGTCCCCGGTCCGCGGATATGGGGGACTCGAGGTTTCGATCAGCCGAA
14161 -----+-----+-----+-----+-----+-----+ 14220
GGCCGTCGGGTGTCCAGGGGCCAGGCGCCTATACCCCTGAGCTCCAAGCTAGTCGGCTT
9-* * G F -

GGTCAGAGCCACGTGGCCGAGGTTCGAGCCCGGAGTTGCCGGCGCCGAGGTTACAGGCGGC
14221 -----+-----+-----+-----+-----+-----+ 14280
CCAGTCTCGGTGCACCGGCTCCAGCTCGGGCCTCAACGGCCGCGGCTCCAATGTCCGCCG
9 T L A V H G L D L G S N G A G L N C A A -

CGTGGCGCAGTCGACGCTGCCGACCGGCGTGCCTTCGGGCGTGGAGCCCGTGTACGACTT
14281 -----+-----+-----+-----+-----+-----+ 14340
GCACCGCGTCAGCTGCGACGGCTGGCCGCACGGAAGCCCGCACCTCGGGCACATGCTGAA
9 T A C D V S G V P T G E P T S G T Y S K -

GCGCACGACGAAGCTGAACGACGCCGCTCCGGACGCGTCCGTGGTGAAGGACGTTCGCGGT
14341 -----+-----+-----+-----+-----+-----+ 14400
CGCGTGCTGCTTCGACTTGCTGCGGCGAGGCCTGCGCAGGCACCACTTCCTGCAGCGCCA
9 R V V F S F S A A G S A D T T F S T A T -

CGCCGGGTTGCACGCGTCTTGCCACCGACCGGAGCGCACTGGGCGATGTAGTAGGTCTC
14401 -----+-----+-----+-----+-----+-----+ 14460
GCGGCCCAACGTGCGCAGGACCGGTGGCTGGCCTCGCGTGACCCGCTACATCATCCAGAG
9 A P N C A D Q G G V P A C Q A I Y Y T E -

GCCGGCGGCGGCACCGCTGACCGACACCGACACGCTCTGTCCGTCACTCAGACCCGAGGC

```

```

14461 -----+-----+-----+-----+-----+-----+ 14520
CGGCCCGCCCGCTGGCGACTGGCTGTGGCTGTGCGAGACAGGCAGTGAGTCTGGGCTCCG
9      G A A A G S V S V S V S Q G D S L G S A -

GGGACTGACGGAGAAGGCGGGCGCGCGCAAGGCGACGGACTGTGCGGCGGCGGCCAGGCC
14521 -----+-----+-----+-----+-----+-----+ 14580
CCCTGACTGCCTCTTCCGCCCGCGCCGCTTCCGCTGCCTGACACGCCGCCGCGGTCCGG
9      P S V S F A P A A F A V S Q A A A A L G -

GATGGATGCGACGGCCACGACGCCGAACCTGGAAGCACGGCGGGACATGTGACGTAACGA
14581 -----+-----+-----+-----+-----+-----+ 14640
CTACCTACGCTGCCGGTGCTGCGGCTTGGACCTTCGTGCCGCCCTGTACACTGCATTGCT
9      I S A V A V V G F R S A R R S M H R L S -

CATGCGTAGGCTCCGATTTCGAGGAGGGGGTTGATCACTCCATGAAAGGATCACCTCGCCG
14641 -----+-----+-----+-----+-----+-----+ 14700
GTACGCATCCGAGGCTAAGCTCCTCCCCCAACTAGTGAGGTACTTTCCTAGTGAGCGGC
9-<      M -
8-*                                     * R A -

GACGGCCCGCCTGCATCTCCCTCTGTGCTCTCGTGGATTTCGGGCACGGCACTCCCGTCCA
14701 -----+-----+-----+-----+-----+-----+ 14760
CTGCCGGCGGACGTAGAGGGAGACACGAGAGCACCTAAAGGCCGTGCCGTGAGGGCAGCT
8      P R G G A D G E T S E H I E P V A S G D -

CGGCCCGCCCGCAGAATGCGGCAGACCCCCCGCACCTCCTCCGGCCCCACCGCCGTACCGG
14761 -----+-----+-----+-----+-----+-----+ 14820
GCCGGCGGGCGTCTTACGCCGTCTGGGGGGCGTGGAGGAGGCCGGGGTGGCGGCATGGCC
8      V A A R L I R C V G R V E E P G V A T G -

TGGGCAGCGACAGCACCCGCTCGGTGAGCGCCTCCACCTTCGGGAGCGGATCGGGCGCGT
14821 -----+-----+-----+-----+-----+-----+ 14880
ACCCGTCGCTGTCTGGGCGAGCCACTCGCGGAGGTGGAAGCCCTCGCCTAGCCCCGCGCA
8      T P L S L V R E T L A E V K P L P D P A -

GGCGCGCGAGGTTCGGACCGGTAGGGCTCGCAGCTGTGGCAGCCGGGGCTGAAGTAGGCGC
14881 -----+-----+-----+-----+-----+-----+ 14940
CCGCGCGCTCCAGCCTGGCCATCCCAGCGCTCGACACCGTCGGCCCCGACTTCATCCGCG
8      H R A L D S R Y P E C S H C G P S F Y A -

GGGCCAGGACGTTGTGCCGTTGGAGCACCGCCTGGAGTTCGTGCGGTCGACCCCGGCGC
14941 -----+-----+-----+-----+-----+-----+ 15000
CCCGGTCTGCAACACGGCAACCTCGTGGCGGACCTCAAGCAGCGCCACGTCGGGCCGCG
8      R A L V N H R Q L V A Q L E D R H L G A -

GGACGGCGTCCACCTCGATGACGACGTACTGGCAGTTCGACAGCTCGTTCGGATCCTGCG
15001 -----+-----+-----+-----+-----+-----+ 15060
CCTGCCGCGAGGTGGAGCTACTGCTGCATGACCGTCAAGCTGTGAGCAAGCCTAGGACGC
8      R V A D V E I V V Y Q C N S L E N P D Q -

GGCGGACCCGGACGCCGGGCAGTCCGTCGAGGTACTGCTCGTACAGACGGTAGTTGCGCC
15061 -----+-----+-----+-----+-----+-----+ 15120
CCGCCCTGGGCCTGCGGCCCGTCAGGCAGCTCCATGACGAGCATGTCTGCCATCAACGCGG
8      P R V R V G P L G D L Y Q E Y L R Y N R -

GGTTGATCGCGGTGAAGTGATCGGCGGACTCCAGGGAGGTGAGGCCCATGGCCGCGCTGA
15121 -----+-----+-----+-----+-----+-----+ 15180
CCAAGTAGCGCCACTTCACTAGCCGCTGAGGTCCCTCCACTCCGGGTACCGGCGCGACT
8      R N I A T F H D A S E L S T L G M A A S -

TCTCGTGCATCCGCGCGACCGTTCCGCTCCCGGTGATCTCATGCGCGGCGTTGAGCCCCT
15181 -----+-----+-----+-----+-----+-----+ 15240
AGAGCACGTAGGCGCGCTGGCAAGGCGAGGGCCACTAGAGTACGCGCCGCAACTCGGGGA
8      I E H M R A V T G S G T I E H A A N L G -

GGTGGCGCATGGCCCGGAGCCGGTCCGCCAGGGCGTCTCGTTCGGTGACGATCGCCCCGC

```


6 A P D V G A V R M G A A R F M A L Q T G -
 ACGTTGCTGCCCAGCTCCACGACCAGGTCGCCGAGGCGAGGCTTGCCCGGCGGGTCGCC
 17581 -----+-----+-----+-----+-----+ 17640
 TGCAACGACGGGTCGAGGTGCTGGTCCAGCGGCTCCGCTCCGAACGGGCGCCAGCGG
 6 V N S G L E V V L D G S A L S A R R T A -
 AGCCCCACGATGTGCGCCATGTGCTCGCGGATCTGGTCGGAGTCGGAGGAGACGTAGACG
 17641 -----+-----+-----+-----+-----+ 17700
 TCGGGCTGCTACACGCGGTACACGAGCGCCTAGACCAGCCTCAGCCTCCTCTGCATCTGC
 6 L G V I H A M H E R I Q D S D S S V Y V -
 TAGTGCTTGAACAGTGTCCCGGGTTCGACGACATGGCGAAGCGTCATCAGCCGGCACGAC
 17701 -----+-----+-----+-----+-----+ 17760
 ATCACGAACCTTGTACAGGGCCCCAGCTGCTGTACCGCTTCGAGTAGTCGGCCGTGCTG
 6 Y H K F L T G P D V V H R L T M L R C S -
 CGGCACACGATGACGTCGAGCGGGAAGACGTCCTGCGCCTCATCGGCGTCGGCCGGATCG
 17761 -----+-----+-----+-----+-----+ 17820
 GCCGTGTGCTACTGCAGCTCGCCCTTCTGCAGGACGCGGAGTAGCCGAGCCGGCCTAGC
 6 R C V I V D L P F V D Q A E D A D A P D -
 ACGAACCCGTTGGCCAGCGGCAGCGAGCCGAAGGAGATCACCTCGGTCCAGTCGTCCGCA
 17821 -----+-----+-----+-----+-----+ 17880
 TGCTTGGGCAACCGGTCGCCGTCGCTCGGCTTCTCTAGTGAGCCAGGTCAGCAGGCGT
 6 V F G N A L P L S G F S I V E T W D D A -
 CCGCATACACGGCACGTCTCGTCCCGCCTGCATTCTCCAGCATGAAGTCTCCTGACGGC
 17881 -----+-----+-----+-----+-----+ 17940
 GGCGTATGTGCCGTGCAGAGCAGGGCGGACGTAAAGAGGTCGTACTTCAGAGGACTGCCG
 6-< G C V R C T E D R R C K E L M -
 GAATGCCGACGCATCGGGCCCCGTCGGTCCGGGGACGGTCAATCTAGGGTTCGGGCCGACG
 17941 -----+-----+-----+-----+-----+ 18000
 CTTACGGCTGCGTAGCCCGGGCAGCCAGGCCCCCTGCCAGTTAGATCCCAAGGCCGGCTGC
 GCGCTCCACTTCGTATGTGCCCTACTGGTTCAGCGGAGCGGACGGGTGAACGCCCCGTAC
 18001 -----+-----+-----+-----+-----+ 18060
 CCGCGAGGTGAAGCATACAGGGATGACCAAGTCGCCTCGCCTGCCACTTGCGGGCATG
 17-* * R L P R T F A R V -
 GTCCTCGATGAGGAGCTGCGGCTGCTCCATGGCCGCGAAGTGCCCGCCGCGGTGGAATC
 18061 -----+-----+-----+-----+-----+ 18120
 CAGGAGCTACTCCTCGACGCCGACGAGGTACCGGCGCTTCACGGGCGGCGCCAGCTTGAG
 17 D E I L L Q P Q E M A A F H G G R D F E -
 GGTCCACCGCGTCAGGGTCGGCAGGATGCCCTCGGCGAACGACCGGATCGGCCGGGTGGC
 18121 -----+-----+-----+-----+-----+ 18180
 CCAGGTGGCGCAGTCCCAGCCGTCCTACGGGAGCCGCTTGCTGGCCTAGCCGGCCCCACCG
 17 T W R T L T P L I G E A F S R I P R T A -
 GTCGTCCGGGAACACCGCGACGCCGACGGGGGCGTTCAGCGGCCAGGGCCCCGCCCCAGGT
 18181 -----+-----+-----+-----+-----+ 18240
 CAGCAGGCCCTTGTGGCGCTGCGGCTGCCCCCGGACGTCGCCGTCCCGGGCGGGTCCA
 17 D D P F V A V G V P A T L P W P G G W T -
 GCGGGCGAAGTCCGCCATGCCGCGAGCCGACTCGTAGTACAACCTGAGCGCTGGAACCGGC
 18241 -----+-----+-----+-----+-----+ 18300
 CGCCCGCTTCAGGCGGTACGGCGCTCGGCTGAGCATCATGTTGACTCGCGACCTTGCCCG
 17 R A F D A M G R A S E Y Y L Q A S S G A -
 CGTCGCGGTTCAGCCAGTAGATCATCACGTGGGTGAGCAGCCGGTCCCGGGAGATGGCCTC
 18301 -----+-----+-----+-----+-----+ 18360
 GCAGCGCCAGTCGGTCATCTAGTAGTGCACCCACTCGTCGGCCAGGGCCCTCTACCGGAG
 17 T A T L W Y I M V H T L L R D R S I A E -

18361
17
18420
18421
17
18480
18481
17
18540
18541
17
18600
18601
17
18660
18661
17
18720
18721
17
18780
18781
17
18840
18841
17
18900
18901
17
18960
18961
17
19020
19021
17
19080
19081
17
19140

CTCCACGTTCTTGGCCGCCGCTCCACTCCTTGGAACTTGTCGAGAATCCAGGCGAGCTGGCC
GAGGTGCAAGAACGGCGGCGAGGTGAGGACCTTGAACAGCTCTTAGGTCCGCTCGACCGG
E V N K G G S W E Q F K D L I W A L Q G -
GACCGGGGAGTTCGGTGAGGCCGTAGGCCAGGGTCTGCGGGCGGGTGGCCTGGATGCGCTG
CTGGCCCCCTCAGCCACTCCGGCATCCGGTCCCAGACGCCCGCCACCGGACCTACGCGAC
V P S D T L G Y A L T Q P R T A Q I R Q -
CCAGCCGATGCCGGTGTTCGGCGAACTCCCCGCTGTGCGCCAGCTTGCCAGGTTCGCTCTC
GGTTCGGCTACGGCCACAGCCGCTTGAGGGGCGACACGCGGTGCAACGGGTCCAGCGAGAG
W G I G T D A F E G S H A L K G L D S E -
GTCCAGGCGCCCCGATGGCCTCCGGGGCGTCTTGGGGCGGGAAGGTCACCAGCATGTTTCAG
CAGGTCCGCGGGCTACCGGAGGCCCGCAGGACCCCGCCCTTCCAGTGGTCGTACAAGTC
D L R G I A E P A D Q P P F T V L M N L -
GTGGACGCCGCCACGTGCTCGGGGTTCGGCCAGCCCCAGCTCCAGCGAGACGACCTTTCC
CACCTGCGGCCGGTGCACGAGCCCCAGCCGGTTCGGGGTCGAGGTTCGCTCTGCTGGAAAGG
H V G A V H E P D A L G L E L S V V K G -
CCAGTCGCCGCCCTGGGCGACGTAACGCTTCGTAGCCGAGGCGGTTTCATCAGCTCCGCCCCA
GGTCAGCGGCGGGACCCGCTGCATTGCGAGCATCGGCTCCGCCAAGTAGTCGAGGCGGGT
W D G G Q A V Y R E Y G L R N M L E A W -
GGCGCGTGCATCCGCCGCACGTCCCAGCCCGGCTCGGCAGTCGGGCGGAGAAGCCGTA
CCGCGCACGCTAGGCGCGTGCAGGGTCGGGCCGAGCCGTCAGCCCGGCCCTCTTCGGCAT
A R A I R R V D W G P E A T P G S F G Y -
GCCCCGCATGGAGGGGACGACGACGTGGAAGGCGTCCGCCGGGTTCGCCGCCGTGCGCGCG
CGGGCCGTACCTCCCCTGCTGCTGCACCTTCCGAGGCGGCCAGCGGCGGCACGCGCGC
G P M S P V V V H F A D A P D G G H A R -
CGGGTCGCTCAGCGGCCGATGACGTCGAGGAACTCGGCGACCGAGCCCGGCCAGCCGTG
GCCAGCGAGTCGCCGGGCTACTGCAGCTCCTTGAGCCGCTGGCTCGGGCCGGTTCGGCAC
P D S L P G I V D L F E A V S G P W G H -
GGTGAGGATCAGCGGGATCGCGTCCGGCTCGGGCGAACGCACGTGAAGGAAGTGCACGTC
CCACTCCTAGTCGCCCTAGCGCAGGCCGAGCCCGCTTGCCTGCACTTCCTTACGTGCAG
T L I L P I A D P E P S R V H L F H V D -
GGCGCCGTCGATCGTGGTGACGAACTGGGGGAACCGGTTTCAGCTCGGCCTCCGCGGCACG
CCGCGGCAGCTAGCACCCTGCTTGACCCCTTGCGCAAGTCGAGCCGGAGGCGCCGTGC
A G D I T T V F Q P F A N L E A E A A R -
CCAGTCGTAGCCGTGGCGCCAGTGGTTCGGTTCGCTCCTTGAGGTAGGACAGCGGCACTCC
GGTCAGCATCGGCACCGCGGTACCCAGCCACTCGAGGAACCTCATCCTGTGCGCGTGAGG
W D Y G H R W H D T L E K L Y S L P V G -
GCGGTCCCATCCGGATCCGGGTATCTCGGACCGCCACCGGGTTCGCGTCGATCCGCCGGGT
CGCCAGGGTAGGCCCTAGGCCCATAGAGCCTGCCGGTGGGCCAGCGCAGCTAGGCGGCCCA
R D W G S G P I E S P W R T A D I R R T -
TAAGGTCGTGCAATGTTCGACTGGGTTCGATCTCGATACGGAAGGGACGCACAGTGAATCC

19141 -----+-----+-----+-----+-----+ 19200
 ATTCCAGCAGCTTACAGCCTGACCCAGCTAGAGCTATGCCTTCCCTGCGTGTCACTTAGG
 17-< L T T S H R V P D I E I R F P R M -
 ACCCTCGTGATTGTGGGAGCGGGGCGGCGGAGGCGGCGCCCGATGTGATCCGGGGAC
 19201 -----+-----+-----+-----+-----+ 19260
 TGGGAGACTAACACCCTCGCCCCGCGCTCCGCGGCGGGGCTACACTAGGCCCTG
 CGTGTCTCAGGCCGGTTCGGCCGCGCGGCGCGCCTTCCCGTGCAGGAGAAGGACCGCAC
 19261 -----+-----+-----+-----+-----+ 19320
 GCACAGAGTCCGGCCAAGCCGCGCGCGCGGCGGAAGGGCACGCCTCTTCTGGCGTG
 16-* * A P E A P A A A G E R A S F S R V -
 GGAGGACAGGAAGTTGCGGATCATCGGCATGCCGTGTTCCGGTCCGGAAGCTCTCCGGATG
 19321 -----+-----+-----+-----+-----+ 19380
 CCTCCTGTCTTCAACGCCTAGTAGCCGTACGGCACAAGCCAGGCCTTCGAGAGGCCTAC
 16 S S L F N R I M P M G H E T R F S E P H -
 GAACTGGACGGACTCCACCGGCAGCGAACGGTGGCGCAGGCCCATCACGTACCCGTCGTC
 19381 -----+-----+-----+-----+-----+ 19440
 CTTGACCTGCCTGAGGTGGCCGTGCTTGCCACCGCGTCCGGGTAGTGCATGGGCAGCAG
 16 F Q V S E V P L S R H R L G M V Y G D D -
 CGTGGAGCGCCCGGTGACCTCGAGGGACGGCGGGACCGTGCCCTCCGGCACGATCAGTGA
 19441 -----+-----+-----+-----+-----+ 19500
 GCACCTCGCGGGCCACTGGAGCTCCCTGCCGCCCTGGCACGGGAGGCCGTGCTAGTCACT
 16 T S R G T V E L S P P V T G E P V I L S -
 GTGGTAGCGGGTCGCGAAGAACCCCGCGGGCAGCCCGGTGAACACTCCGCGCCCGTCGTG
 19501 -----+-----+-----+-----+-----+ 19560
 CACCATCGCCCAGCGCTTCTTGGGGCGCCCGTCGGGCCACTTGTGAGGCGCGGGCAGCAC
 16 H Y R T A F F G A P L G T F V G R G D H -
 CGTGATCCGGCTCGTCTTCCCGTGCATGAGATGCCGGGCGGGGACGGTGGCGGCGCCGTA
 19561 -----+-----+-----+-----+-----+ 19620
 GCACTAGGCCGAGCAGAAGGGCACGTACTCTACGGCCCGCCCTGCCACCGCCGCGGCAT
 16 T I R S T K G H M L H R A P V T A A G Y -
 GGCGCGGGCGACGGCCTGATGCCCCAGACAGACCCCGAGCAGCGGGACCCGGCCGGCGAA
 19621 -----+-----+-----+-----+-----+ 19680
 CCGCGCCCGCTGCCGGAATACGGGGTCTGTCTGGGGCTCGTCGCCCTGGGCGCGCCGCTT
 16 A R A V A Q H G L C V G L L P V R G A F -
 GGCCTGGACGATCTCGACGTGCCCCGAGGTGTGCGGGTGGCCGGGGCCCGGCCCCAGCAG
 19681 -----+-----+-----+-----+-----+ 19740
 CCGGACCTGCTAGAGCTGCACGGGCCTCCACAGCCCCACCGCCCCGGGCGGGTCTGTC
 16 A Q V I E V H G S T D P H G P G P G L L -
 GACCGCGTCCGGCCGCATCAGCCCCATCTCGTCCGGGGTCATGAGATGCGACCGCACCAT
 19741 -----+-----+-----+-----+-----+ 19800
 CTGGCGCAGGCCGGCGTAGTCGGGGTAGAGCAGGCCCCAGTACTCTACGCTGGCGTGGTA
 16 V A D P R M L G M E D P T M L H S R V M -
 GACGGGCTCCGCGCCGGCGGACATCAGATACTGGCGCAGGATGTCGACGAAGCTGTGAA
 19801 -----+-----+-----+-----+-----+ 19860
 CTGCCCCGAGGCGCGGCCGCTGTAGTCTATGACCGCGTCTACAGCTGCTTCGACAGCTT
 16 V P E A G A S M L Y Q R L I D V F S D F -
 CGCGTCGACCACCAGGACCCGCGGGGCTCGGTGCCTGCGCCGATCCGTGCGGAGACCA
 19861 -----+-----+-----+-----+-----+ 19920
 GCGCAGCTGGTGGTCTTGGGCGCCCCGAGCCACGGACGCGGCCTAGGCAGCCCTCTGGT
 16 A D V V L V R P A E T G A G S G D P S W -
 CAAGCTCACAGCAACTCCTCTCCGGTGACCGCCAGTGAGTGGCGCTCATCTTGGCCAGC
 19921 -----+-----+-----+-----+-----+ 19980
 GTTCGAGTGTGTTGAGGAGAGGCCACTGGCGGGTCACTCACCGCGAGTAGAACCGGTG

15 R E V D N A F V E P A S D D T V A S A P -
AAGCCGCTCTGCGGACGGATCGGGCCAGGCCCTGCTCCACGTCGCGCAGGAGCCCGGTGACG
20761 -----+-----+-----+-----+-----+-----+-----+ 20820
TTCGGCAGACGCCTGCCTAGCCCCGGTCCGGACGAGGTGCAGCCGCTCCTCGGGCCACTGC
15 F G D A S P D P W A Q E V D A L L G T V -
GTCTCCGGCGCGAGGCCGGGCCAGTACGGGGACTCGTGGAGCAGCAGTTTCGCATCGGCCG
20821 -----+-----+-----+-----+-----+-----+-----+ 20880
CAGAGGCCGCGCTCCGGCCCCGGTCATGCCCTGAGCACCTCGTCGTCAAGCGTAGCCGGC
15 T E P A L G P W Y P S E H L L L E C R G -
GTGGCGAGATCGGTGACCACGCTGCCCCGGTGCAGGACCATGCGTACGTCCGGCAGGCCA
20881 -----+-----+-----+-----+-----+-----+-----+ 20940
CACCGCTCTAGCCACTGGTGCAGCGGGGCCACGTCCTGGTACGCATGCAGGCCGTCCGGT
15 T A L D T V V S G R H L V M R V D P L G -
GGCCGGTTCTCGATGAGGTGGGGCAGGTCTCTCGATGTAGCGGGCCGTGTCGTACCCGAAG
20941 -----+-----+-----+-----+-----+-----+-----+ 21000
CCGGCCAAGAGCTACTCCACCCCGTCCAGGAGCTACATCGCCCGGCACAGCATGGGCTTC
15 P R N E I L H P L D E I Y R A T D Y G F -
AACCCGAGGAACCCGAAGCGGAAGCCGGACGCGGACCCCTCGGCGTCGAACATGTCCCGC
21001 -----+-----+-----+-----+-----+-----+-----+ 21060
TTGGGCTCCTTGGGCTTCGCCTTCGGCCTGCGCCTGGGGAGCCGAGCTTGTACAGGGCG
15 F G L F G F R F G S A S G E A D F M D R -
ATGGCCCGCAGCAGCGGCCACAACCCGCCCGCGGTACGCAGCCGAGCCCCCTGGGGGCCG
21061 -----+-----+-----+-----+-----+-----+-----+ 21120
TACGGGCGTCTGTCGCCGGTGTTGGGCGGGCGCCATGCGTCGGCGTCGGGGACCCCCGGC
15 M A R L L P W L G G A T R L R L G Q P G -
TCCTCCAGGAGCGCGCCGGCCCGCTCCAGGAGCAGGCCCCCGCAGGGCGGGTACGCCCTCG
21121 -----+-----+-----+-----+-----+-----+-----+ 21180
AGGAGGTCTCTCGCGCGGCCGGGCGAGGTCTCTCGTCCGGGGCGTCCCGCCCATGCGGGAGC
15 D E L L A G A R E L L L G R L A P V G E -
ACGCGCACCAACCCGGTCGGTGACCGAGAGCGAGAGCAGCGCGCCGAAGCCGACGAAGTGG
21181 -----+-----+-----+-----+-----+-----+-----+ 21240
TGCGCGTGGTGGGCCAGCCACTGGCTCTCGCTCTCGTCGCGCGGCTTCGGCTGCTTGACC
15 V R V V R D T V S L S L L A G F G V F Q -
TGCCTGCGGTGCGGGGCCGGGCCGGCCGCGGACTCCAGGAGGTAGACCTCGTCGGGGCCG
21241 -----+-----+-----+-----+-----+-----+-----+ 21300
ACGGACGCCAGCGCCCGGCCGGCCGGCGCCTGAGGTCTCCATCTGGAGCAGCCCCGGC
15 H R R D R A P G A A S E L L Y V E D P G -
AAGTGCTCGGCCAGCGCGCGGTAGGCGGGCAGGGCGCCCGTCTCCTTCACATCGAGGCGT
21301 -----+-----+-----+-----+-----+-----+-----+ 21360
TTCACGAGCCGGTCGCGCGCCATCCGCCCCGTCCCGCGGGCAGAGGAAGTGTAGCTCCGCA
15 F H E A L A R Y A P L A G T E K V D L R -
CGTGTCGCGCACCCCGCACCGGGGCCGAGACCACGCACTGGTCGGTCATCCTGGGTCTCCC
21361 -----+-----+-----+-----+-----+-----+-----+ 21420
GCACAGGCGTGGGCGTGGCCCCGGCTCTGGTGCCTGACCAGCCAGTAGGACCCAGGAGGG
15-< R T R V R V P A S V V C Q D T M -
GGATCACGTGGTGATGGCGTAGCGGTGTGCCACCTGACGGGCGGTGACGACCGCCCGGTG
21421 -----+-----+-----+-----+-----+-----+-----+ 21480
CCTAGTGCACACTACCGCATCGCCACACGGTGGACTGCCCCGAGTCGTGGCGGGGCCAG
14-* * T T I A Y R H A V Q R A T L V A R D -
GGGGCCGGAGCGGTGTGTCGACGACGCGCGCGGCCCTTCCAGCTGACGAAGGAGCCGGTGTG
21481 -----+-----+-----+-----+-----+-----+-----+ 21540
CCCCGGCTCGCCAACAGCTGCTGCGCGCGCCGGAAGGTGCACTGCTTCCTCGGCCACAC
14 P G S R N D V V R A A K W S V F S G T H -

```

GGTCACGGGGTCGAGGTCGGTGTCCACGACGATGCCGGCGTGCGCGCCGGTCCGCTCCCT
21541 -----+-----+-----+-----+-----+-----+ 21600
CCAGTGCCCCAGCTCCAGCCACAGGTGCTGTACGGCCGCACGCGCGGCCAGGCGAGGGA
14      T V P D L D T D V V I G A H A G T R E R -

GAGCCGGGCGGCGACGGCCTCGCCGATGCCCTGCCGTTCCTCCCTCGGCGCCGGCCAGCAG
21601 -----+-----+-----+-----+-----+-----+ 21660
CTCGGCCCCGCGCTGCCGAGCGGCTACGGGACGGCAAGGGGGAGCCGCGGCCGGTCTGTC
14      L R A A V A E G I G Q R E G E A G A L L -

GTCCATGCGCACGGTGACGGCGTCTGCTGCCGTCTCTGCCGGTTCGATGACGACCTGGTA
21661 -----+-----+-----+-----+-----+-----+ 21720
CAGGTACGCGTGCCACTGCCGACGCGACGGCAGCAGGACGGCCAGCTACTGCTGGACCAT
14      D M R V T V A D S G D D Q R D I V V Q Y -

GCCGAGGCAGCCGCCGACCCCGTCGAGGATCGCGGCCCTCCAGCTCGGCGGGCTGGAGGGT
21721 -----+-----+-----+-----+-----+-----+ 21780
CGGCTCCGTTCGGCGGCTGGGGCAGCTCCTAGCGCCGGAGGTCGAGCCGCCCGACCTCCCA
14      G L C G G V G D L I A A E L E A P Q L T -

CACGTGCGCCAGGGGGATGCGGTCCGCGACCCGGCCGATGACCTGGATCCGCGGTCCCGG
21781 -----+-----+-----+-----+-----+-----+ 21840
GTGCAGCGGGTCCCCCTACGCCAGGCGCTGGGCGGCTACTGGACCTAGGCGCCAGGGCC
14      V D G L P I R D A V R G I V Q I R P G P -

CAGCGGCTCCCCGGGGCCCCCGGGAGGATGCGGACCAGGTCCCCGGTTCGGTAGCGGAT
21841 -----+-----+-----+-----+-----+-----+ 21900
GTCGCCGAGGGGGCCCCGGGCGGCCCTCTACGCTGGTCCAGGGGCCACGCCATCGCCTA
14      L P E G P G A P L I R V L D G T R Y R I -

CAGTGGTTTGTATGCCGTCCACCAGCATGGTGAGGACGAGTTCGCCCTCTCCCGTGTGCC
21901 -----+-----+-----+-----+-----+-----+ 21960
GTCACCAAATACGGCAGGTGGTCTGTAACCTCTGCTCAAGCGGGAGAGGGCACAGCGG
14      L P K I G D V L M T L V L E G E G T D G -

GACCACGGCGCCGGTGTCCGGTTCGACGAGTTCGGTCAAGTAGTTGGGCTGGGCGAGGTG
21961 -----+-----+-----+-----+-----+-----+ 22020
CTGGTGCCGCGGCCACAGGCCAAGCTGCTCAAGCCAGTTCATCAACCCGACCCGCTCCAC
14      V V A G T D P E V L E T L Y N P Q A L H -

GAGCGCTCCGGTGTCCGCTCCGGTGGCGATGCACAGGGCTTCCTGGGAGCCGTAGAGCGT
22021 -----+-----+-----+-----+-----+-----+ 22080
CTCGCGAGGGCCACAGGCGAGGCCACCGCTACGTGTCCCGAAGGACCCTCGGCATCTCGCA
14      L A G T D A G T A I C L A E Q S G Y L T -

GGGCCGCACGACGGCTTGCGGCCAGAGGGTCGCCACGTTGTGCGCGAACTGCGGGGTGCA
22081 -----+-----+-----+-----+-----+-----+ 22140
CCCGGCGTGTGCGGAACGCGGTCTCCAGCGGTGCAACAGCCGCTTGACGCCCCACGT
14      P R V V A Q P W L T A V N D A F Q P T C -

GATCTCACCCAGCGTGAGGAAGAGCTTCACGGGAAGCCGGGCCAGGTCGTAGCCGTAGTG
22141 -----+-----+-----+-----+-----+-----+ 22200
CTAGAGTGGGTGCGACTCCTTCTCGAAGTGCCCTTCGGCCCGGTCCAGCATCGGCATCAC
14      I E G L T L F L K V P L R A L D Y G Y H -

CAGGGCCGCCTTGGAAGGCTCAGGCACAGCGCCGGAGCACAGACGACCTCGACCTC
22201 -----+-----+-----+-----+-----+-----+ 22260
GTCCCGGCGGAACCGTTCCGAGTCCGTGTGCGGCCCTCGTGTCTGCTGGAGCTGGAG
14      L A A K A L S L C L A P A C V V V E V E -

CAGCTCCTCGATCAGCCGACGCGCTTACGGAATCCCACCCTGGGGGACTCGGGCCAGAT
22261 -----+-----+-----+-----+-----+-----+ 22320
GTCGAGGAGCTAGTCGGCGTTCGCGGAATGCCTTAGGGTGGGACCCCTGAGCCCGGTCTA
14      L E E I L R L A K R F G V R P S E P W I -

```

[illegible]

23101 AGGATCATCCGGTTGAGCAGGGCATTGACGGTCAGCTGAGCCCATACTCGCCGGCGCTG 23160
 21 -----+-----+-----+-----+-----+-----+-----+
 TCCTAGTAGGCCAACTCGTCCCCTAACTGCCAGTCGACTCGGGTATGGAGCGGCCGCGAC
 L I M R N L L A N V T L Q A W V E G A S -
 23161 TAGCGGCGGGCGACCGAGATGATCCCCGCGACCTTGTGCTCAGCGGCCGGTCTGAAGCGC 23220
 21 -----+-----+-----+-----+-----+-----+-----+
 ATCGCCGCCCGCTGGCTCTACTAGGGGCGCTGGAACAACGAGTCGCCGGCCAGCTTCGCG
 Y R R A V S I I G A V K N S L P R D F R -
 23221 AGATAACCGACTCCGGCACGCTCGATGAAGGTCTGCATGAGGCTGGCCGTGCCGAATCCG 23280
 21 -----+-----+-----+-----+-----+-----+-----+
 TCTATTGGCTGAGGCCGTGCGAGCTACTTCCAGACGTAAGTCCGACCGGCACGGCTTAGGC
 L Y G V G A R E I F T Q M L S A T G F G -
 23281 TGCACGGGCGCCGCGAAGATGATCCCGTCCGCGCGACCATCTTCGCCACGACCTCGGGC 23340
 21 -----+-----+-----+-----+-----+-----+-----+
 ACGTGCCCGCGGCGCTTCTACTAGGGCAGGCGGCGCTGGTAGAAGCGGTGCTGGAGCCCG
 H V P A A F I I G D A A V M K A V V E P -
 23341 ACCCGTTCGGCCAGGGTGCAGGCCACCGGCCTGTCTGTTGCAGTCCCCGAGGGCCCCGAC 23400
 21 -----+-----+-----+-----+-----+-----+-----+
 V G D A L T C A V P R D N C D G C P G C -
 23401 CGCTCCATCCTGATCGAGCGCAGGTTCGACGGCCTCGAAGTCGACGCCGCGGTTCTCTGCT 23460
 21 -----+-----+-----+-----+-----+-----+-----+
 GCGAGGTAGGACTAGCTCGCGTCCAGCTGCCGGAGCTTCAGCTGCGGCGCCAAGAGACGA
 R E M R I S R L D V A E F D V G R N E A -
 23461 ACGCGTGCCGCGTGCCGCGAGTACGTGCGCGGTGTTGCCGTACGTTCCGAACCGTTGATC 23520
 21 -----+-----+-----+-----+-----+-----+-----+
 TGCGCACGGCGCACGGCGTTCATGCAGCGGCCACAACGGCAGTGCAAGGCTTGGAACCTAG
 V R A A H R L V D A T N G D R E S G N I -
 23521 GCGAGGATCTTGAGTTGTGCGCTCACGAGGGGCTCCTTGGTGAGTCAGGTGCGCTCGGC 23580
 13-* -----+-----+-----+-----+-----+-----+-----+
 21-< CGCTCCTAGAACTCAACACGCGAGTGCTCCCCGAGGAACCACTCAGTCCACGCGAGCCG
 * T R E A -
 A L I K L Q A S M -
 23581 GGTTCGGCTCGGGGAACTGTCTGGCCGCCGCTGGTCCGGGAGCCGAGGGCCGGCTCGGC 23640
 13 -----+-----+-----+-----+-----+-----+-----+
 CCAGCCGAGCCCCCTTGACAGACCGGCGGCGACAGGCCCTCGGCGTCCCGGCCGAGCCG
 T P E P S S D P R R Q D P L R L A P E A -
 23641 GGGGGCGGGAGGAAGACCGCCCCGCGGCGGGCCGCCACGCTCGCCGAACCGGATGAGGGG 23700
 13 -----+-----+-----+-----+-----+-----+-----+
 CCCCCGCCCTCCTTCTGGCGGGGCGCGCGGCGGTGCGAGCGGCTTGGCCTACTCCCC
 P A P P L G G R P P G G R E G F R I L P -
 23701 CTTCTCGACGAGATAGAAGCTGATGGTCGCCAGCACGACGCTGATCGAGATCGTGAAGAG 23760
 13 -----+-----+-----+-----+-----+-----+-----+
 GAAGAGCTGCTCTATCTTCGACTACCAGCGGTGCTGCTGCGACTAGCTCTAGCACTTCTC
 K E V L Y F S I T A L V V S I S I T F L -
 23761 GAACAGTTCACGAAACCCATGTACCCCCGGAATTCCGGCGTTGGCACGGGAGACTTGCC 23820
 13 -----+-----+-----+-----+-----+-----+-----+
 CTTGTCAAGGGTCTTGGGGTACAGTGGGGCCTTAAGGCCGCAACCGTGCCCTCTGAACGG
 F L E W F G M D G R F E P T P V P S K G -
 23821 GAAGATGCTGCCGTTCTCTGAGCCAGAGGTTGATCACGATCTCGTGCCAGAGGTAGACGCC 23880
 13 -----+-----+-----+-----+-----+-----+-----+
 CTTCTACGACGGCAAGGACTCGGTCTCCAAGTGTAGAGCACGGTCTCCATCTGCGG
 F I S G N R L W L N I V I E H W L Y V G -
 GAGGGAGATCTGGCCGAGGAAGAGGATCGGCTTGCTGGTGAAGAGCGCGTCCGAGAACCG

```

23881 -----+-----+-----+-----+-----+-----+-----+ 23940
CTCCCTCTAGACCGGCTCCTTCTCCTAGCCGAACGACCACCTTCTCGCGCAGGCTCTTGGC
13      L S I Q G L F L I P K S T F L A D S F R -

GGACTCGGCGCCGGGGACCGTCATCGGTGCCAGGAGCAGCAGGGTGAAGGAGGTCAAGGAT
23941 -----+-----+-----+-----+-----+-----+-----+ 24000
CCTGAGCCGCGGGCCCTGGCAGTAGCCACGGTCCTCGTCGCCACTTCCTCCAGTCCTA
13      S E A G P V T M P A L L L L T F S T L I -

GAAGTGGTCGACGAGCTCCTGGGCCAGGGCCGCGTTGTGCGCCCATGCCCGGGATGCCGAT
24001 -----+-----+-----+-----+-----+-----+-----+ 24060
CTTCACCAGCTGCTCGAGGACCCGGTCCCGGCGCAACAGCGGTACGGGCCCTACGGCTA
13      F H D V L E Q A L A A N D G M G P I G I -

GGGCTTGGTGGCGTAGAGGAGGTACAGCGGGATGAGCGGGACCCAGCAGATCAGCGGGCG
24061 -----+-----+-----+-----+-----+-----+-----+ 24120
CCCGAACCACCGCATCTCCTCCATGTGCGCCCTACTCGCCCTGGGTCTGTCTAGTCGCCCCG
13      P K T A Y L L Y L P I L P V W C I L P R -

CCGGATCACGAAACGGTAGAAGCCCGGGGTCCCTGGCGTCGCCTCGGCGTACGCGGAGTA
24121 -----+-----+-----+-----+-----+-----+-----+ 24180
GGCCTAGTGCTTTGCCATCTTCGGGCCCCAGGGACCGCAGCGGAGCCGCATGCGCCTCAT
13      R I V F R Y F G P T G P T A E A Y A S Y -

GATGGCCAGTGCCATGCCCGCGGCGAAGCAGCCGGCGTAGTAGGGCGGCCAGTACCACTG
24181 -----+-----+-----+-----+-----+-----+-----+ 24240
CTACCGGTACGGTACGGGCGCGCTTCGTGCGCCGCATCATCCCGCCGGTCAAGGTGAC
13      I A L A M G A A F C G A Y Y P P W Y W Q -

CATCGTCGCGCCGGTGGAGGGGAGGTTGGTGTACGTGACCCAGCCGATGGCCATGACTTC
24241 -----+-----+-----+-----+-----+-----+-----+ 24300
GTAGCAGCGCGGCCACCTCCCTCCAACCATGCACTGGGTGCGGTACCGGTACTGAAG
13      M T A G T S P L N T Y T V W G I A M V E -

CAGCGCGGCCAGCGGCAGCAGGAGGCGGCGTGCCTTCTGCCCGGGAGTGCTGCCGCCCCG
24301 -----+-----+-----+-----+-----+-----+-----+ 24360
GTCGCGCCGGTCGCGCTCGTCTCCGCGCACGGAAGACGGGCGCCTCACGACGGCGGGGC
13      L A A L P L L L R R A K Q G P T S G G R -

CGCGAGCCGGTGGCCGATCCAGGCGATCAGCGGCAGGGCGAGGTAGAACGTGAACCTCGGC
24361 -----+-----+-----+-----+-----+-----+-----+ 24420
GCGCTCGGCCACCGGCTAGGTCCGCTAGTCGCGCTCCCGCTCCATCTTGCACCTGAGCCG
13      A L R H G I W A I L P L A L Y F T F E A -

GGGGACCGTCCAGGTGGGCTCGATGCCGTGCATCGGCTGGCCCTCGGGCAGATAGAAGTG
24421 -----+-----+-----+-----+-----+-----+-----+ 24480
CCCCTGGCAGGTCCACCCGAGCTACGGCACGTAGCCGACCGGGAGCCCGTCTATCTTCAC
13      P V T W T P E I G H M P Q G E P L Y F H -

CATGAGCAGCACGGGCCGAGGACGTGCTGACGCTGTGATCTCGAACCAGTTGTAGCC
24481 -----+-----+-----+-----+-----+-----+-----+ 24540
GTACTCGTCTGCGCCGCGTCTGTCAGCGACTGCGACAGCTAGAGCTTGGTCAACATCGG
13      M L L V P R L V D S V S D I E F W N Y G -

GGGGATTGCGAAGACGAGCAACAGGTAGTAGGCGGGCAGGATGCGCAGGGCCCCGGCGTTT
24541 -----+-----+-----+-----+-----+-----+-----+ 24600
CCCCTAACGCTTCTGCTCGTTGTCCATCATCCGCGCGTCTACGCGTCCCGGGCCGCAAA
13      P I A F V L L L Y Y A P L I R L A R R K -

GAGGAACCGTCCGGTGGCGGGCGCTTCGTCCCACTGATGGTGACGCGGGCGTAGGGCTT
24601 -----+-----+-----+-----+-----+-----+-----+ 24660
CTCCTTGGCAGGCCACCGCCCGGCGAAGCAGGGTGACTACCACTGCGCCCGCATCCCGAA
13      L F R G T A P R K T G S I T V R A Y P K -

GTACAGCATCATTCCGGACAGAGCGAAGAAGGGGAAGGCATACCCCCAGACCGTCCCGG
24661 -----+-----+-----+-----+-----+-----+-----+ 24720

```

13-< CATGTCGTAGTAAGGCCTGTCTCGCTTCTTCCCCCTTCCGTATGGGGGTCTGGCAGGCGC
 Y L M M G S L A F F P S P M -
 24721 AGGACGCCCCAGAACGGTTTGGCCGGCTCACCGACGAAGCTGCCCCACTCCGGCCTGGAAG
 -----+-----+-----+-----+-----+-----+ 24780
 TCCTGCGGGGTCTTGCCAAACGGGCCGAGTGGCTGCTTCGACGGGTGAGGCCGACCTTC
 GCGACGTGGTAGACGACCACACCCAGCGCGAGGACACCTCGCAGTCCCTCGAACTTCGGT
 24781 -----+-----+-----+-----+-----+-----+ 24840
 CGCTGCACCATCTGCTGGTGTGGGTGCGGCTCCTGTGGAGCGTCAGGGAGCTTGAAGCCA
 ATTCGCTTGCTTTTTGCGCCACCTGCGTCGCGAAGGACGTCCCCATGGAACAGTCCCCCT
 24841 -----+-----+-----+-----+-----+-----+ 24900
 TAAGCGAACGAAAAACGCGGTGGACGCAGCGCTTCCTGCAGGGGTACCTTGTGAGGGGA
 TTCCCTTGGCACTTGCTCGTTGACTTCCCGAAATAGTCGGGTCTGCGGAGTGTGAGCCGC
 24901 -----+-----+-----+-----+-----+-----+ 24960
 AAGGGAACCGTGAACGAGCAACTGAAGGGCTTTATCAGCCCAGACGCCTCACACTCGGCG
 ATCTCCAATCGTGCTGTTCCGGTGCTCAGGACGACTTGTTCGGCCTGAGTGGGAAGGCA
 24961 -----+-----+-----+-----+-----+-----+ 25020
 TAGAGGTTAGCACGACAAGGCCACGAGTCCTGCTGAACAAAGCCGACTCACCTTCCGT
 12-* * S S K N R G S H S P -
 GCCACCCCCGCGCCCCGCTCGGCCAGACCGGGGGCCGAGGAGTCCCGTTCGAGAGGA
 25021 -----+-----+-----+-----+-----+-----+ 25080
 CCGTGCGGGGCGGCGGGGCGGAGCCGGTCTGGCCCCCGGCTCCTCAGGGCAAGGCTCTCCT
 12 L W G R R G A E A L G P A S S D R E S L -
 TCGGAGTGATCTCCGGCGGCCAGGCGATGCCACCTCCGGATCCAGCGGATTCAAGCCAT
 25081 -----+-----+-----+-----+-----+-----+ 25140
 AGCCTCACTAGAGGCCGCGGTCCGCTACGGGTGGAGGCCTAGGTGCGCTAAGTTCGGTA
 12 I P T I E P P W A I G V E P D L P N L G -
 GTTCGAGCCGCGGGGTGCTAGGCCGCGGAGCACAGGTAGACGATCACCGCCTCGTCGCTCA
 25141 -----+-----+-----+-----+-----+-----+ 25200
 CAAGCTCGGCCCCCAGCATCCGGCGGCTCGTGTCCATCTGCTAGTGGCGGAGCAGCGAGT
 12 H E L R P D Y A A S C L Y V I V A E D S -
 GCGTGAGGAATCCGAAGCCCAGCCCCGCGGAGACGTACAGCGCCCGTCCGTTCTCCTCGC
 25201 -----+-----+-----+-----+-----+-----+ 25260
 CGCACTCCTTAGGCTTCGGGTGCGGGCGCCTCTGCATGTCGCGGGCAGGCAAGAGGAGCG
 12 L T L F G F G L G A S V Y L A R G N E E -
 CGAGCTCCACGGTCCGCCAGCCGCGCAAGGTGGGCGACCCACCCGGATGTCGACCACGG
 25261 -----+-----+-----+-----+-----+-----+ 25320
 GCTCGAGGTGCCAGGCGGTGCGGCGCTTCCACCCGCTGGGGTGGGCCTACAGCTGGTGCC
 12 G L E V T R W G G F T P S G V R I D V V -
 CGCCGAACACGCTGCCGCGCAGGCAGCTGAAGTACTTGGCCTGGCCGGGTACGCCCCCGG
 25321 -----+-----+-----+-----+-----+-----+ 25380
 GCGGCTTGTCGACGGCGCGTCCGTCGACTTCATGAACCGGACCGGCCCATGCGGGGGCC
 12 A G F V S G R L C S F Y K A Q G P V G G -
 CGAAGTGGATGCCCCGAGCACCCCGTGGGAGGAGATCGCGCAGTTCGCCTGCCGAGGT
 25381 -----+-----+-----+-----+-----+-----+ 25440
 GCTTCACCTACGGGGCGTCGTGGGGCACCTCCTCTAGCGCGTCAAGCGGACGGCGTCCA
 12 A F H I G R L V G H S S I A C N A Q R L -
 CGAAGGAGTGGCCTACGGTGCGGCGGAAGGGCTCGCCCTGGAACCACTCGCGAAACGAGC
 25441 -----+-----+-----+-----+-----+-----+ 25500
 GCTTCCTCACCGGATGCCACGCCGCTTCCCGAGCGGGACCTTGGTGAGCGCTTTGCTCG
 12 D F S H G V T R R F P E G Q F W E R F S -
 CCCGTTGCTACGGAAGACCTGCTTCTCCTCCGTCCACGCTCCCGAGATCCCGATCGGCT
 25501 -----+-----+-----+-----+-----+-----+ 25560


```

12      GGGCAAGCAGTGCCTTCTGGACGAAGAGGAGGCAGGTGCGAGGGCTCTAGGGCTAGCCGA
      G R E D R F V Q K E E T W A G S I G I P -
      TCATCGCTGGCCCCCTTCTCTCGACTTCTCTCGACGACTCGCGGGAGGCGGCCGAGGGGTC
25561  -----+-----+-----+-----+-----+-----+ 25620
      AGTAGCGACCGGGGAAGAGAGCTGAAGAGAGCTGCTGAGCGCCCTCCGCCGGCTCCCCAG
12-<    K M -
      CGCCGGGCCCCGTGGGAACGCCGCGAGTCTAGATGCGGCGGCACCGGGGGCAGGGGGGTGCG
25621  -----+-----+-----+-----+-----+-----+ 25680
      GCGGCCCCGGGCACCCTTGCGGCGTCAGATCTACGCCGCGGTGGCCCCCGTCCCCCACGC
      GACGACGTCCGCCCCACCTCAGCACACCGGGAGATGCAGGTGCGGTGACGGGCGACGTGAC
25681  -----+-----+-----+-----+-----+-----+ 25740
      CTGCTGCAGGCGGGGTGGAGTCGTGTGGCCCTCTACGTCCAGCCACTGCCCGCTGCACTG
      GATGCAACGGTCCGAGGCCCCGTTGCCCCGACGACGGCCACAGAGCCATCGGAGCAACG
25741  -----+-----+-----+-----+-----+-----+ 25800
      CTACGTTGCCAGGCTCCGGGCCAACGGGCCTGCTGCCGGGTGTCTCGGTAGCCTCGTTGC
      GAGGCGGACCGCAGATGACCAAGCACGCCCGTGACCGCGCGGTAGTCCTCGGCGCAGGGA
25801  -----+-----+-----+-----+-----+-----+ 25860
      CTCCGCCTGGCGTCTACTGGTTTCGTGCGGGCACTGGCGCGCCATCAGGAGCCGCGTCCCT
20->          M T K H A R D R A V V L G A G M -
      TGGCGGGGCTGCTCGCCGCGCGCTCTGTCCGAGACGTACAAGGAAGTGTGGTGTATCG
25861  -----+-----+-----+-----+-----+-----+ 25920
      ACCGCCCCGACGAGCGGCGCGCGCAGGACAGGCTCTGCATGTTCTTTCACGACCACTAGC
20      A G L L A A R V L S E T Y K E V L V I D -
      ACCGGGACCGGTTGGGCGGCACGGAGCAGCGCCGCGGTGTCCCGCACGGACGCCACGCCC
25921  -----+-----+-----+-----+-----+-----+ 25980
      TGGCCCTGGCCAACCCGCGCTGCCTCGTTCGCGGCGCCACAGGGCGTGCCTGCGGTGCGGG
20      R D R L G G T E Q R R G V P H G R H A H -
      ATGCGCTGCTGGCCAAGGGACAGCAGATCCTCAACGAACCTTTCCCCGGACTCGACACCG
25981  -----+-----+-----+-----+-----+-----+ 26040
      TACGCGACGACCGGTTCCCTGTGCTCTAGGAGTTGCTTGAGAAGGGGCGTGCCTGTGGC
20      A L L A K G Q Q I L N E L F P G L D T E -
      AACTCACCTCGGCCGGAATCCCCGCGGGGACATCGCCGGGAACCTGCGGTGGTACTTCA
26041  -----+-----+-----+-----+-----+-----+ 26100
      TTGAGTGGAGCCGGCCTTAGGGGCGGCCCTGTAGCGGCCCTTGGACGCCACCATGAAGT
20      L T S A G I P A G D I A G N L R W Y F N -
      ACGGCCGCGGCTCCAGCCCTTCGACACCGGGCTGATCAGCGTCTCGGCGACGAGGCCCG
26101  -----+-----+-----+-----+-----+-----+ 26160
      TGCCGGCGGCGAGGTGCGGAAGCTGTGGCCCGACTAGTCGCAGAGCCGCTGCTCCGGGC
20      G R R L Q P F D T G L I S V S A T R P E -
      AGCTGGAGTCCCACGTGCGCGCACGGGTGCGCGCTGCCACAGGTGAAGATCATGGACG
26161  -----+-----+-----+-----+-----+-----+ 26220
      TCGACCTCAGGGTGACGCGCGTGGCCAGCGGCGGACGGTGTCCACTTCTAGTACCTGC
20      L E S H V R A R V A A L P Q V K I M D G -
      GGTGCGTGATCCGGGGCCTGACCGCCTCGGCCGACCGCAGCCGCGTCACCGGTGTGAGG
26221  -----+-----+-----+-----+-----+-----+ 26280
      CCACGCACTAGGCCCCGACTGGCGGAGCCGGCTGGCGTGGCGCAGTGGCCACAGCTCC
20      C V I R G L T A S A D R S R V T G V E V -
      TGGTCGACGAGTGGGTACGGACACCCCGACGCGCCTGGAGGCGGACCTCGTGTGACG
26281  -----+-----+-----+-----+-----+-----+ 26340
      ACCAGCTGCTCAGCCCATGCCTGTGGGGCTGCGCGGACCTCCGGCTGGAGCAGCAGCTGC
20      V D E S G T D T P T R L E A D L V V D V -
      TCACGGGGCGCGGCTCGCGGACTCCCGCCTGGCTGGAGGAGTTCGGATACGAGCGGCCCG

```

26341 -----+-----+-----+-----+-----+-----+-----+ 26400
 AGTGTCCCGCGCCGAGCGCCTGAGGGCGGACCGACCTCCTCAAGCCTATGCTCGCCGGGC
 20 T G R G S R T P A W L E E F G Y E R P A -
 CGGAGGACCGCTTCAAGATCGATCTGGCGTACACCACGCGCCACTTCAAGCTCAAGGAAG
 26401 -----+-----+-----+-----+-----+-----+ 26460
 GCCTCCTGGCGAAGTTCTAGCTAGACCGCATGTGGTGC GCGGTGAAGTTCGAGTTCCTTC
 20 E D R F K I D L A Y T T R H F K L K E D -
 ACCCCTACGGCACGGACCTGTGATCAACCCGGTGGCATCGCCGAGCAACCCGCGCGGCG
 26461 -----+-----+-----+-----+-----+-----+ 26520
 TGGGGATGCCGTGCCTGGACAGCTAGTTGGGCCACCGTAGCGGCTCGTTGGGCGCGCCGC
 20 P Y G T D L S I N P V A S P S N P R G A -
 CGTTCTTCCCCCGGCTCGCGGACGGCAGCTCCCAGCTCTCCCTCACCGGAATCCTCGGCG
 26521 -----+-----+-----+-----+-----+-----+ 26580
 GCAAGAAGGGGGCCGAGCGCCTGCCGTGAGGGTCGAGAGGGAGTGGCCTTAGGAGCCGC
 20 F F P R L A D G S S Q L S L T G I L G D -
 ACCACCCGCCCACCGACGACGAGGGCTTCTGGCGTTCGCCAAGTCGCTTGCCGCGCCGG
 26581 -----+-----+-----+-----+-----+-----+ 26640
 TGGTGGGCGGGTGGCTGCTGCTCCCGAAGGACCGCAAGCGGTTGAGCGAACGGCGCGGCC
 20 H P P T D D E G F L A F A K S L A A P E -
 AGATCTACCGGGCCGTCCGCGATGCCGAACCTCTCGACGAACCGGTCACCTTCCGCTTCC
 26641 -----+-----+-----+-----+-----+-----+ 26700
 TCTAGATGGCCCGGCAGGCGCTACGGCTTGGAGAGCTGCTTGGCCAGTGGAAGGCGAAGG
 20 I Y R A V R D A E P L D E P V T F R F P -
 CGGCGAGCGTCCGCGCCCGTTACGAGAGGCTGCGCCGTTTCCCCGGCGGGTTCTCTGTC
 26701 -----+-----+-----+-----+-----+-----+ 26760
 GCCGCTCGCAGGCGGCGGCAATGCTCTCCGACGCGGCAAAGGGGCGCCCAAGGAGCAGT
 20 A S V R R R Y E R L R R F P G G F L V M -
 TGGGCGACGGCGTGTGCAGCTTCAACCCCGTCTACGGCCAGGGCATGACGGTCGCCGCCC
 26761 -----+-----+-----+-----+-----+-----+ 26820
 ACCCGCTGCCGCACACGTGCAAGTTGGGGCAGATGCCGGTCCCGTACTGCCAGCGGCGGG
 20 G D G V C S F N P V Y G Q G M T V A A L -
 TGGAGGCCGTGGCGCTGCGGGACCACTTGC GCGACGCCCCGACCCCGACGCCCTGCGCT
 26821 -----+-----+-----+-----+-----+-----+ 26880
 ACCTCCGGCACCGCGACGCCCTGGTGAACGCGCTGCGGGGCTGGGGCTGCGGGACGCGA
 20 E A V A L R D H L R D A P D P D A L R F -
 TCTTCCGGCGTATCTCCACGGTCATCGACGTTCCGTGGGACATCGCCGCCGAGCGGATC
 26881 -----+-----+-----+-----+-----+-----+ 26940
 AGAAGGCCGCATAGAGGTGCCAGTAGCTGCAAGGCACCCTGTAGCGGCGGCCTCGCCTAG
 20 F R R I S T V I D V P W D I A A G A D L -
 TGAACCTCCCCGGGGTGGAGGGCCCCCGCACCATGAAGGTGAAGATGGCCAACGCCTACA
 26941 -----+-----+-----+-----+-----+-----+ 27000
 ACTTGAAGGGGCCCCACCTCCCGGGGGCGTGGTACTTCCACTTCTACCGGTTGCGGATGT
 20 N F P G V E G P R T M K V K M A N A Y M -
 TGGCCCGCTGCACGCAGCGGCAGCCGTGACGCGCGGTTGACCGGGGCGTTCTTCCGGG
 27001 -----+-----+-----+-----+-----+-----+ 27060
 ACCGGGCGGACGTGCGTGC GCGTCCGCGAGCTGCCGCGCCACTGGCCCCGCAAGAAGGCCC
 20 A R L H A A A A V D G A V T G A F F R V -
 TGGCCGGGCTGGTGGACCCCCCGCAGGCCCTGATGCGCCCCCTCCCTCGCCCTGCGGGTCA
 27061 -----+-----+-----+-----+-----+-----+ 27120
 ACCGGCCCCGACCACCTGGGGGGCGTCCGGGACTACGCGGGGAGGGAGCGGGACGCCAGT
 20 A G L V D P P Q A L M R P S L A L R V M -
 TGCGCAACTCCTCGGCGAAGCCGTGCGTCCCTTCGGGCGCCCGCTATGACCGCGCGGCC
 27121 -----+-----+-----+-----+-----+-----+ 27180

ACGCGTTGAGGAGCCGCTTCGGCAGCCAGGGAAGCCCGCGGCGGCATACTGGCGCGCCGG
 20-* R N S S A K P S V P S G A A V * -

CGTCCGGGGCGGCTGCCGGGGCCAGGAGCCGACATGCGGGTGATGATCACGGTGTTCCCG
 27181 -----+-----+-----+-----+-----+ 27240
 GCAGGCCCCCGCGACGGCCCCGGTCCTCGGCTGTACGCCCCTACTAGTGCCACAAGGGC
 19 M R V M I T V F P -

GCGCGGGCGCACTTCCTGCCGCTGGTGCCCTATGCCTGGGCCCCTGCAGAGCGCGGGCCAC
 27241 -----+-----+-----+-----+-----+ 27300
 CGCGCCCCCGGTGAAGGACGGCGACCACGGGATACGGACCCGGGACGTCTCGCGCCCCGGTG
 19 A R A H F L P L V P Y A W A L Q S A G H -

GAGGTATGTGTCGTGGCGCCCCCGGGCTATCCCACCGGGGTGGCCGACCCCGACTTCCAC
 27301 -----+-----+-----+-----+-----+ 27360
 CTCCATACACAGCACCGCGGGGGCCCGATAGGGTGGCCCCACCGGTGGGGCTGAAGGTG
 19 E V C V V A P P G Y P T G V A D P D F H -

GAGGCCGTCACCGCGGCCCGCCTGAAGTCGGTGACCTGCGGGCAGCCGCAGCCGCTGGCG
 27361 -----+-----+-----+-----+-----+ 27420
 CTCCGGCAGTGGCGCCGGCCGGACTTCAGCCACTGGACGCCCCGTCGGCGTCGGCGACCCG
 19 E A V T A A G L K S V T C G Q P Q P L A -

GTCCACGACCGCGACGACCCCGGCTACGCGGCGATGCTGCCGACCGCGGCGGAGTCGGAG
 27421 -----+-----+-----+-----+-----+ 27480
 CAGGTGCTGGCGCTGCTGGGGCCGATGCGCCGCTACGACGGCTGGCGCCGCTCAGCCTC
 19 V H D R D D P G Y A A M L P T A A E S E -

CGCTACGTGGCGGCCCTCGGGATCAGCGAGAAGGAGCGCCCCACCTGGGACGTCTTCTAC
 27481 -----+-----+-----+-----+-----+ 27540
 GCGATGCACCGCCGGGAGCCCTAGTCGCTCTTCTCGCGGGGTGGACCTGCAGAAGATG
 19 R Y V A A L G I S E K E R P T W D V F Y -

CACTTCACCTTGCTGGCGATCCGCGACTACCATCCGCCGCGGCCGCGGACGAGCTGGAC
 27541 -----+-----+-----+-----+-----+ 27600
 GTGAAGTGGAACGACCGCTAGGCGCTGATGGTAGGCGGCGCCGGCGCCGTCTGCACCTG
 19 H F T L L A I R D Y H P P R P R Q D V D -

CAGGTGATCGAGTTTCGCCCCGATCTGGCAGCCCCGATCTGGTGCTGTGGGACCCCTGGTTC
 27601 -----+-----+-----+-----+-----+ 27660
 GTCCACTAGCTCAAGCGGGCCTAGACCGTCGGGCTAGACCACGACACCCTGCGGACCAAG
 19 Q V I E F A R I W Q P D L V L W D A W F -

CCCTCGGGCGCGATCGCGGCGCGGGTCAGCGGCGCCGCGCACGCGCGGGTGCTCGTAGCC
 27661 -----+-----+-----+-----+-----+ 27720
 GGGAGCCCCGCTAGCGCCGCGCCAGTCGCGCGGCGCGTGCAGCGCCACGAGCATCGG
 19 P S G A I A A R V S G A A H A R V L V A -

CCCCACTACACCGGTGGGTACCGAGCGGTTCCGCCCGCGGGCCCCCGCGGGGGGCC
 27721 -----+-----+-----+-----+-----+ 27780
 GGGCTGATGTGGCCGACCCAGTGGCTCGCCAAGCGGCGGCGCCGGGGCGCCGCCCCCG
 19 P D Y T G W V T E R F A A A G P A A G A -

GACCTCCTGGCCGAGACGATGCGGCCGCTGGCCGAGCGGTACGGCGTGAGGTGACGAC
 27781 -----+-----+-----+-----+-----+ 27840
 CTGGAGGACCGCTCTGCTACGCCGCGACCGGCTCGCCATGCCGCACCTCCAGCTGCTG
 19 D L L A E T M R P L A E R Y G V E V D D -

GATCTTCTGCTCGGACAGTGGACGGTCAATCCGTTCCCGGCGCCGATGAACCCGCCGACC
 27841 -----+-----+-----+-----+-----+ 27900
 CTAGAAGACGAGCCTGTACCTGCCAGTTAGGCAAGGGCCGCGGCTACTTGGGCGGCTGG
 19 D L L L G Q W T V N P F P A P M N P P T -

CGGCTCACGAACGTTCCGGTGCGCTACGTGCCCTACACCGGTGCCAGCGTCATGCCCGCG
 27901 -----+-----+-----+-----+-----+ 27960
 GCCGAGTGCTTGCAAGGCCACGCGATGCACGGATGTGGCCACGGTCGCGAGTACGGGCGC

```

19      R L T N V P V R Y V P Y T G A S V M P A  -
      TGGCTGTACGCGCGGCCGTCGCGGCCGCGGGTGGCGCTGTCGCTCGGAGTGTCCGCGCGG
27961  -----+-----+-----+-----+-----+ 28020
      ACCGACATGCGCGCGGCCGAGCGCCGCGGCCACCGCGACAGCGAGCCTCACAGGCGCGCC
19      W L Y A R P S R P R V A L S L G V S A R  -
      GCGTTCCTCAAGGGTGACTGGGGGCGTACCGCCAACTGCTGGAAGCGGTGCGGAGCTG
28021  -----+-----+-----+-----+-----+ 28080
      CGCAAGGAGTTCCCACTGACCCCCGCATGGCGGTTTGACGACCTTCGCCAGCGCCTCGAC
19      A F L K G D W G R T A K L L E A V A E L  -
      GACATCGAGGTGATCGCCACGCTCAACGACAACCAACTGGCGGAGAGCGGGCCGCTGCCG
28081  -----+-----+-----+-----+-----+ 28140
      CTGTAGCTCCACTAGCGGTGCGAGTTGCTGTTGGTTGACCGCTCTCGCCCGGCGACGGC
19      D I E V I A T L N D N Q L A E S G P L P  -
      GACAACGTCCACACCCTCGACTACGTACCGCTCGACCAGTTGCTGCCCACCTGCTCGGCC
28141  -----+-----+-----+-----+-----+ 28200
      CTGTTGCAGGTGTGGGAGCTGATGCATGGCGAGCTGGTCAACGACGGGTGGACGAGCCGG
19      D N V H T L D Y V P L D Q L L P T C S A  -
      GTCATCCACCACGGATCGACGGGCACCTTCGCCCGGCGAGCGCGCGCGGGCTGCCCCAG
28201  -----+-----+-----+-----+-----+ 28260
      CAGTAGGTGGTGCCTAGCTGCCCGTGAAGCGGCGCCGCTCGCGCCGGCCCGACGGGGTC
19      V I H H G S T G T F A A A S A A G L P Q  -
      GTGGTCTGCGACACCGACGAGCCCCCTCCTGCTCTTCGGCGAGGACACCCCGACGGCATC
28261  -----+-----+-----+-----+-----+ 28320
      CACCAGACGCTGTGGCTGCTCGGGGAGGACGAGAAGCCGCTCCTGTGGGGGCTGCCGTAG
19      V V C D T D E P L L L F G E D T P D G I  -
      GCGTGGGACTTCACCTGCCAGAAGCAGCTCACCGGACGCTCACCTCCCGCGTGGTCACC
28321  -----+-----+-----+-----+-----+ 28380
      CGCACCTTGAAGTGGACGGTCTTCGTCGAGTGGCGCTCGGAGTGGAGGGCGCACCACTGG
19      A W D F T C Q K Q L T A T L T S R V V T  -
      GACTACGGGGCGGGGGTGC GCGTCGACCACCAGAAGCAGTCCGCCGGACAGATCCGTGAG
28381  -----+-----+-----+-----+-----+ 28440
      CTGATGCCCCCGCCCCACGCGCAGCTGGTGGTCTTCGTCAGGCGGCCTGTCTAGGCACTC
19      D Y G A G V R V D H Q K Q S A G Q I R E  -
      CAACTACGCAGGGTGCTCACCGAACCTTCCTTCCGCGAGGGCGCTCGACGGATCCGGGAA
28441  -----+-----+-----+-----+-----+ 28500
      GTTGATGCGTCCCACGAGTGGCTTGAAGGAAGGCGCTCCCGCGAGCTGCCTAGGCCCTT
19      Q L R R R V L T E P S F R E G A R R I R E  -
      GACCGGAATTCCGCCCCCAGCCCGGTGCAACTCGTATCGCTCCTGGTAGAACTGACGAAG
28501  -----+-----+-----+-----+-----+ 28560
      CTGGCCTTAAGGCGGGGGTCCGGCCAGCTTGAGCATAGCGAGGACCATCTTGACTGCTTC
19      D R N S A P S P V E L V S L L V E L T K  -
      CGTCATCGCCGTGACAAGGAGGCGGACCGATGAGGATGCTGGTGACGGGCGGAGCGGGTT
28561  -----+-----+-----+-----+-----+ 28620
      GCAGTAGCGGCACTGTTCTCCGCTGGCTACTCCTACGACCACTGCCCGCCTCGCCCAA
19-*   R H R R D K E A D R * -
1->      M R M L V T G G A G F -
      TCATCGGCTCGCAGTTCGTGCGGGCCACACTGCACGGCGAGCTGCCGGTTCCGAGGACG
28621  -----+-----+-----+-----+-----+ 28680
      AGTAGCCGAGCGTCAAGCACGCCCCGGTGTGACGTGCCGCTCGACGGCCCAAGGCTCCTGC
1      I G S Q F V R A T L H G E L P G S E D A  -
      CCCGGGTGACGGTCTCTGGACAAGCTGACGTACTCCGGCAATCCGGCCAACCTCACCTCCG
28681  -----+-----+-----+-----+-----+ 28740
      GGGCCCACTGCCAGGACCTGTTCTGACTGCATGAGGCCGTTAGGCCGGTTGGAGTGGAGGC

```

1 R V T V L D K L T Y S G N P A N L T S V -

28741 TCGCGGCCCCATCCGCGGTACACCTTCGTCCAGGGCGACACCGTCGACCCGCGCGTTCGTTCG 28800
-----+-----+-----+-----+-----+
AGCGCCGGGTAGGCGCCATGTGGAAGCAGGTCCCCTGTGGCAGCTGGGCGCGCAGCAGC
A A H P R Y T F V Q G D T V D P R V V D -

28801 ACGAGGTGGTCGCGGCGCCACGACGTCATCGTCCACTTCGCGGCGGAGTCGCACGTGGACC 28860
-----+-----+-----+-----+-----+
TGCTCCACCAGCGGCCGGTGTGTCAGTAGCAGGTGAAGCGCCGCCTCAGCGTGCACCTGG
E V V A G H D V I V H F A A E S H V D R -

28861 GCTCGATCGACACCGCCACCCGGTTCGTACGACCAACGTGCTCGGGACCCAGACGCTGC 28920
-----+-----+-----+-----+-----+
CGAGCTAGCTGTGGCGGTGGGCCAAGCAGTGTGTTGCACGAGCCCTGGGTCTGCGACG
S I D T A T R F V T T N V L G T Q T L L -

28921 TGGAAGCGGCTCTCCGGCACGGGTCGGCCGGTTCGTGCACGTGTGCACCGACGAGGTCT 28980
-----+-----+-----+-----+-----+
ACCTTCGCCGAGAGGCCGTGCCCCAGCCGGCCAAGCACGTGCACAGCTGGCTGCTCCAGA
E A A L R H G V G R F V H V S T D E V Y -

28981 ACGGGTCGATCGCCTCCGGCTCATGGACCGAGGACACCCCGCTCGCCCCCAACGTCCCCT 29040
-----+-----+-----+-----+-----+
TGCCCAGCTAGCGGAGGCCGAGTACCTGGCTCCTGTGGGGCGAGCGGGGGTTGCAGGGGA
G S I A S G S W T E D T P L A P N V P Y -

29041 ACGCGGCGTCAAGGCGGGTTTCGGACCTGATGGCGCTCGCCTGGCACCCGACCCGGGGCC 29100
-----+-----+-----+-----+-----+
TGCGCCGCGAGCTTCCGCCCAAGCCTGGACTACCGCGAGCGGACCGTGGCGTGGGCCCCGG
A A S K A G S D L M A L A W H R T R G L -

29101 TGGACGTGTCGTCAACCGGTGCACCAACAACCTACGGTCCCCTACCAGTACCCCGAGAAGG 29160
-----+-----+-----+-----+-----+
ACCTGCAGCAGCAGTGGGCCACGTGGTTGTTGATGCCAGGGATGGTCATGGGGCTCTTCC
D V V V T R C T N N Y G P Y Q Y P E K V -

29161 TGATCCCGCTCTTCGTACCAACATCCTCGACGGCTTGCGGGTGCCCCGTGACGGGGACG 29220
-----+-----+-----+-----+-----+
ACTAGGGCGAGAAGCAGTGGTTGTAGGAGCTGCCGAACGCCACGGGGACATGCCCCCTGC
I P L F V T N I L D G L R V P L Y G D G -

29221 GCGCCACCGCCGGGACTGGCTGCACGTGTCCGACCACTGCCGGGCCATCCAGATGGTCA 29280
-----+-----+-----+-----+-----+
CGCGGGTGGCGGCCCTGACCGACGTGCACAGGCTGGTGACGGCCCGGTAGGTCTACCACT
A H R R D W L H V S D H C R A I Q M V M -

29281 TGAATCCCGGCCGGGCGGGGAGGTCTACCACATCGCGCGGCGGCACCGAACTCTCCAACG 29340
-----+-----+-----+-----+-----+
ACTTGAGGCCGGGCCCGCCCCCTCCAGATGGTGTAGCCGCGCGCGTGGCTTGAGAGGTTGC
N S G R A G E V Y H I G G G T E L S N E -

29341 AGGAACTCACCGGCCTGTTGCTCACGGCGTGCGGCACCGACTGGTCCTGCGTGGACCGGG 29400
-----+-----+-----+-----+-----+
TCCTTGAGTGGCCGACAAACGAGTGCCGCACGCCGTGGCTGACCAGGACGCACCTGGCCCC
E L T G L L L T A C G T D W S C V D R V -

29401 TGGCCGACCGGCAGGGGCACGACCGCCGCTACTCGCTCGACATCACGAAGATCCGGCAGG 29460
-----+-----+-----+-----+-----+
ACCGGCTGGCCGTCCCCTGCTGGCGGCGATGAGCGAGCTGTAGTGCTTCTAGGCCGTCC
A D R Q G H D R R Y S L D I T K I R Q E -

29461 AACTGGGCTACGAGCCCCCTGGTCGCCCTTCGAGGACGGCCTGGCCGCGACGGTGAAGTGGT 29520
-----+-----+-----+-----+-----+
TTGACCCGATGCTCGGGGACACGCGGAAGCTCCTGCGGACCGGCGCTGCCACTTCACCA
L G Y E P L V A F E D G L A A T V K W Y -

AAGGAGCTGGCCCGGACCGGGTGGGACCCGCTCGCCGCGGCGCGGTGGTCCCTCGGCGTG
 30301 -----+-----+-----+-----+-----+ 30360
 TTCCTCGACCGGGCCTGGCCACCCCTGGGCGAGCGGCGGCGCCACCAGGAGCCGCAC
 2 K E L A R T G W D P L A A G A V V L G V -

 ATCTTCGGCGCGCTGTTCTGTCAGCGCCAGCGGCGGTTGGCCGACCCCATGCTGGACCTC
 30361 -----+-----+-----+-----+-----+ 30420
 TAGAAGCCGCGGACAAGCAGGTGCGGGTGGCGCCCAACCGGCTGGGGTACGACCTGGAG
 2 I F G A L F V Q R Q R R L A D P M L D L -

 GGCCTCTTCGCGGACCGCACCCCTGCGGGCGGGTCTGACGGTCAGTCTGGTCAACGCCGTC
 30421 -----+-----+-----+-----+-----+ 30480
 CCGGAGAAGCGGCTGGCGTGGGACGCGCCGCGGAGTCCAGTCCAGACAGTTGCGGCAG
 2 G L F A D R T L R A G L T V S L V N A V -

 ATCATGGGCGGGACCGGACTGATGGTTCGCCCTGTACCTCCAGACGATCGCCGGTCACTCC
 30481 -----+-----+-----+-----+-----+ 30540
 TAGTACCCGCGCTGGCCTGACTACCAGCGGGACATGGAGGTCTGCTAGCGGCCAGTGAGG
 2 I M G G T G L M V A L Y L Q T I A G H S -

 CCGTTGGCCCGCGGGCTGTGGCTGCTGATCCCGGCCTGCATGCTCGTCTGGGCGTACAG
 30541 -----+-----+-----+-----+-----+ 30600
 GGCAACCGGCGGCGCCGACACCGACGACTAGGGCCGGACGTACGAGCAGCACCCGCATGTC
 2 P L A A G L W L L I P A C M L V V G V Q -

 CTGTGCAACCTGCTGGCCAGCGGATGCCCTTCCCGGGTGTCTGGGGGGACTGCTG
 30601 -----+-----+-----+-----+-----+ 30660
 GACAGCTTGGACGACCGGGTGCCTACGGGGGAAGGGCCACGACGACCCCCCTGACGAC
 2 L S N L L A Q R M P P S R V L L G G L L -

 ATCGCGGCGCTCGGACAGCTCCTGATCACCCAGGTGGACACCGAGGACACCGCCCTCCTC
 30661 -----+-----+-----+-----+-----+ 30720
 TAGCGCCGCGAGCCTGTGAGGACTAGTGGGTCCACCTGTGGCTCCTGTGGCGGGAGGAG
 2 I A A V G Q L L I T Q V D T E D T A L L -

 ATCGCGGCCACCACCCTGATCTACTTCGGCGCCTCACCGGTGGGGCCGATCACACGGGC
 30721 -----+-----+-----+-----+-----+ 30780
 TAGCGCCGGTGGTGGGACTAGATGAAGCCGCGGAGTGGCCACCCCGGCTAGTGGTGCCCCG
 2 I A A T T L I Y F G A S P V G P I T T G -

 GCGATCATGGGAGCCGCGCCCCCGGAGAAGGCGGGTGGCGCCTCGTCGCTGTCCGCCACC
 30781 -----+-----+-----+-----+-----+ 30840
 CGCTAGTACCCTCGGCGCGGGGGCCTCTTCCGCCACGGCGGAGCAGCGACAGGCGGTGG
 2 A I M G A A P P E K A G A A S S L S A T -

 GGCGGCGAGTTCGGAGTGGCGCTCGGCATCGCGGGCCTGGGGAGTCTGGGCACCGTCTGTG
 30841 -----+-----+-----+-----+-----+ 30900
 CCGCCGCTCAAGCCTCACCGCGAGCCGTAGCGCCCGGACCCCTCAGACCCGTGGCAGCAC
 2 G G E F G V A L G I A G L G S L G T V V -

 TACAGCGCCGGGGTTCGAGGTGCCGACGCGGCGGGCCCGGACGCGGACGCGCGCAG
 30901 -----+-----+-----+-----+-----+ 30960
 ATGTGCGGGCCCCAGCTCCACGGCCTGCGCCGGCCCGGGCGGCTGCGGGCGCGTC
 2 Y S A G V E V P D A A G P A D A D A A Q -

 GAGAGCATCGCCGGCGCCCTGCACACGGCCGGTTCAGTGGCACCGGGCAGCGCCGACGCC
 30961 -----+-----+-----+-----+-----+ 31020
 CTCTCGTAGCGGCCGCGGGACGTGTGCCGGCCAGTCGACCGTGGCCCGTTCGCGGCTGCGG
 2 E S I A G A L H T A G Q L A P G S A D A -

 CTGCTGGACTCCGCGCGCGCGGCCTTACCAGCGGCGTGCAGTCCGTGCGCCGCCGTCTGC
 31021 -----+-----+-----+-----+-----+ 31080
 GACGACCTGAGGCGCGCGCGCGGAAGTGGTTCGCCGACGTCAGGCAGCGGCGGCAGACG
 2 L L D S A R A A F T S G V Q S V A A V C -

```

31081  GCGGTGTTCTCCCTGGCGCTCGCCGTCCTCATCGGCACCCGGCTGCGGGACATTTCCGCG
-----+-----+-----+-----+-----+-----+-----+-----+ 31140
CGGCACAAGAGGGACCGCGAGCGGCAGGAGTAGCCGTGGGCGGACGCCCTGTAAAGGCGC
2    A V F S L A L A V L I G T R L R D I S A -

31141  ATGGACCACGGGCACGGCGAGGAACCGGCCGAGAACGACGCTCAACCGGCCACATGAGCG
-----+-----+-----+-----+-----+-----+-----+ 31200
TACCTGGTGCCCGTGCCGCTCCTTGCCCGGCTCTTGCTGCGAGTTGGCCGGTGTACTCGC
2-*  M D H G H G E E P A E N D A Q P A T * -

31201  CACTTCCGGAGATGCAACGGCCGCGCTCGAGGTATGAGGATCACCTTCCGGGGTGACCT
-----+-----+-----+-----+-----+-----+-----+ 31260
GTGAAGGCCTCTACGTTGCCGGCGGCAGCTCCATACTCCTAGTGGAAGCCCCACGTGGA

31261  GCACGGCAACGGAGGCGTAGTGGAGTACTGGAACAGCACGGCGGAGACCATGCCCCGCCA
-----+-----+-----+-----+-----+-----+-----+ 31320
CGTGCCGTTGCCTCCGCATCACCTCATGACCTTGTCGTGCCGCTCTGGTACGGGGCGGT
3->      M E Y W N S T A E T M P R Q -

31321  GGAACTCGAACAGTGGAAGTGGCGCAGGCTCCAGGCCGCCATGGACCACGCCAGAAGGCT
-----+-----+-----+-----+-----+-----+-----+ 31380
CCTTGAGCTTGTACCTTCACCGCTCCGAGGTCCGGCGGTACCTGGTGCGGTCTTCCGA
3    E L E Q W K W R R L Q A A M D H A R R L -

31381  TTCGCCCTTCTGGCGGGAACGACTCCCCGAGAACATCACCTCCATGGCGGACTACGCGGC
-----+-----+-----+-----+-----+-----+-----+ 31440
AAGCGGGAAGACCGCCCTTGCTGAGGGGCTCTTGTTAGTGGAGGTACCGCCTGATGCGCCG
3    S P F W R E R L P E N I T S M A D Y A A -

31441  GCGGGTGCTCTCCTGCGCAAGGCCGACCTCCTCGCCGCGGAAGCCGCTCTCCCCCTTA
-----+-----+-----+-----+-----+-----+-----+ 31500
CGCCACGGAGAGGACGCGTTCGGGCTGGAGGAGCGGCGCCTTCGGCGCAGAGGGGGAAT
3    R V P L L R K A D L L A A E A A S P P Y -

31501  CGGCACCTGGCCCTCGCTGGATCCGGCGCTCGGAGTGCGCCATCACCAGACCAGCGGCAC
-----+-----+-----+-----+-----+-----+-----+ 31560
GCCGTGGACCGGGAGCGACCTAGGCCGCGAGCCTCACGCGGTAGTGGTCTGGTGCCTGTG
3    G T W P S L D P A L G V R H H Q T S G T -

31561  CAGCGGTAACCCCCCATCCGGACGTTGACACCGAACGCGACTGGGCCTGGTGCGTGGA
-----+-----+-----+-----+-----+-----+-----+ 31620
GTGCCATTGGGGGGGTAGGCCTGCAAGCTGTGGCTTGCGCTGACCCGGACCACGCACCT
3    S G N P P I R T F D T E R D W A W C V D -

31621  CACGTTCTGCACGGCGCTCCACAGCATGGGCGTGCGCCCGCACCACAAGGGTCTGGTGGC
-----+-----+-----+-----+-----+-----+-----+ 31680
GTGCAAGACGTGCCGCGAGGTGTGCTACCCGCACGCGGGCGTGGTGTTCACAGACCACCG
3    T F C T A L H S M G V R P H H K G L V A -

31681  GTTCGGCTACGGGCTGTTGCGCCGTTTCTGGGGCATGCACTACGGCCTCGAGCGCATGGG
-----+-----+-----+-----+-----+-----+-----+ 31740
CAAGCCGATGCCCCGACAAGCGGCCAAAGACCCCGTACGTGATGCCGGAGCTCGCGTACCC
3    F G Y G L F A G F W G M H Y G L E R M G -

31741  CGCCACGGTCATCCCGGCCGGCGGCCTCGACTCCCGCTCCCGGGTACGGCTGCTGGTCTGA
-----+-----+-----+-----+-----+-----+-----+ 31800
GCGGTGCCAGTAGGGCCGGCCGCGGAGCTGAGGGCGAGGGCCCATGCCGACGACGACT
3    A T V I P A G G L D S R S R V R L L V D -

31801  CTACCAGATCGAGGTGCTCGGCCTCACACCGAGCTATGCGATGCGGCTGATCGAGACGGC
-----+-----+-----+-----+-----+-----+-----+ 31860
GATGGTCTAGCTCCACGAGCCGAGTGTGGCTCGATACGCTACGCCGACTAGCTCTGCCG
3    Y Q I E V L G L T P S Y A M R L I E T A -

31861  CCGCGAGATGGGCATCGACCTCGCCCGCGAGGCTAACGTCCAGATCATCTGGCCGGGGC
-----+-----+-----+-----+-----+-----+-----+ 31920

```


3 GGCGCTCTACCCGTAGCTGGAGCGGGCGCTCCGATTGCAGGTCTAGTAGGACCGGCCCCG
 R E M G I D L A R E A N V Q I I L A G A -
 31921 GGAGCCGCGCTCCGCGTTTACCACCCGCACCATCGAGGAGGCCTTCGGCGCCCGGGTCTT
 -----+-----+-----+-----+-----+-----+ 31980
 3 CCTCGGCGCGAGGCGCAAGTGGTGGGCGTGGTAGCTCCTCCGGAAGCCGCGGGCCCAGAA
 E P R S A F T T R T I E E A F G A R V F -
 31981 CAACGCCGCGGGCACCCTGAGTTTCGGGGGGGTGTTTCATGTTTCGAGTGCACCGCCCGGCG
 -----+-----+-----+-----+-----+-----+ 32040
 3 GTTTCGGCGCCCGTGGTGAAGCCCCCACAAGTACAAGCTCACGTGGCGGGCCGC
 N A A G T T E F G G V F M F E C T A R R -
 32041 CGAGGCCTGCCACATCATCGAACCCTCGTGCATCGAGGAGGTGCTCGACCCGGTGACGGA
 -----+-----+-----+-----+-----+-----+ 32100
 3 GCTCCGACCGTGTAGTAGCTTGGGAGCAGTAGCTCCTCCACGAGCTGGGCCACTGCCT
 E A C H I I E P S C I E E V L D P V T E -
 32101 ACAGCCCGTTCGGCTACGGCGAGGAGGGCGTCCGAGTCACCACCGGGCTGAACCGTGAGGG
 -----+-----+-----+-----+-----+-----+ 32160
 3 TGTGGGCGAGCCGATGCCGCTCCTCCCGCAGGCTCAGTGGTGGCCCCGACTTGGCACTCCC
 Q P V G Y G E E G V R V T T G L N R E G -
 32161 GATGCAGCTCTTCCGGCACTGGACCGAGGACGTCGTGGTCAAGCGGCCCCACACCGAGTG
 -----+-----+-----+-----+-----+-----+ 32220
 3 CTACGTCGAGAAGGCCGTGACCTGGCTCCTGCAGCACCAGTTCGCCGGGGTGTGGCTCAC
 M Q L F R H W T E D V V V K R P H T E C -
 32221 CGGCTGCGGCGGACGTTGGGACTTCTACGACGGCGGCATCCTTCGGCGCGTGGACGACAT
 -----+-----+-----+-----+-----+-----+ 32280
 3 GCCGACGCCGGCTGCACCCTGAAGATGCTGCCGCCGTAGGAAGCCGCGACCTGCTGTA
 G C G R T W D F Y D G G I L R R V D D M -
 32281 GCGCAAGATACGCGGGGTCTCGATCACCCCGGTGATGATCGAGGATGTGCTGCGCGGCTT
 -----+-----+-----+-----+-----+-----+ 32340
 3 CGCGTTCTATGCGCCCCAGAGCTAGTGGGGCCACTACTAGCTCCTACACGACGCGCCGTA
 R K I R G V S I T P V M I E D V L R G F -
 32341 CGACGAGGTGAACGAGTTCCACTCGTCCATCCGACCGTCCGCGGACTCGATACGATCCA
 -----+-----+-----+-----+-----+-----+ 32400
 3 GCTGCTCCACTTGCTCAAGGTGAGCAGGTAGGCCTGGCAGGCGCCTGAGCTATGCTAGGT
 D E V N E F H S S I R T V R G L D T I H -
 32401 CGTCAAGGTCGAGGCGGGGAGACATCTCGGGTGAAGCGGCCGAGAGCCTGTGCGGCCGCAT
 -----+-----+-----+-----+-----+-----+ 32460
 3 GCAGTTCAGCTCCGCCCTCTGTAGAGCCCACTCCGCCGGCTCTCGGACACGCGGCGTA
 V K V E A G D I S G E A A E S L C G R I -
 32461 CACCGAGGAGTTCAAGCGTGAGATAGGCATACGGCCCCAGGTGGAGCTGACCCCCGCGGG
 -----+-----+-----+-----+-----+-----+ 32520
 3 GTGGCTCCTCAAGTTCGCACTCTATCCGTATGCCGGGGTCCACCTCGACTGGGGGCGCCC
 T E E F K R E I G I R P Q V E L T P A G -
 32521 CAGCCTCCCCCGATCGAAGTGGAAGGCGGCACGACTTCATGACGAGCGGAACTCGCCCC
 -----+-----+-----+-----+-----+-----+ 32580
 3 GTCGGAGGGGGTAGCTTACCTTCCGCCGTGCTGAAGTACTGCTCGCGCTTGAGCGGGG
 S L P R S K W K A A R L H D E R E L A P -
 32581 TCAGGCCTGAGCAGGTGGAGCAGCTCCTGGTGAAGTACCAGGAGCCTGGGCCTGCTGGAGC
 -----+-----+-----+-----+-----+-----+ 32640
 3-* AGTCCGGAAGTCCACCTCGTCGAGGACCACTCGATGGCCTCGGACCCGGACGACCTCG
 Q A * -
 32641 AGAGCTGCGCGGTCCCGGCCGTGCTCGCCGCGGTACAGGGCCGCCCGTGCAGAACTCCGTA
 -----+-----+-----+-----+-----+-----+ 32700
 TCTCGACGCGCCAGGGCCGGCACGAGCGGCGCCAGTCCCGCGGGCACGCCTTGAGGCAT

TCGCCCTGGACGGCCAGGGCGTGGAGTTCGAGTACTACCGGGGCACGACGACGACGACCTCG
32701 -----+-----+-----+-----+-----+ 32760
AGCGGGACCTGCCGGTCCCGCACCTCAAGCTCATGATGGCCCCCGTGCTGCTGTCGGAGC

TGGCCTGAACCCACCCCCGGTCCGCCGGGTCAGACGAAAGGGAGACCGGTGCCCCACGGT
32761 -----+-----+-----+-----+-----+ 32820
ACCGGACTTGGGTGGGGGCCAGGCGGCCAGTCTGCTTTCCTCTGGCCACGGGGTGCCA
4-> M P H G -

GCAGAGCGCGAAGCGAGCCCGGCCGAGGAGAGCGCCGGCACCCGGCCGCTGACCGGCGAG
32821 -----+-----+-----+-----+-----+ 32880
CGTCTCGCGCTTCGCTCGGGCCGGCTCCTCTCGCGGCCGTGGGCCGGCGACTGGCCGCTC
4 A E R E A S P A E E S A G T R P L T G E -

GAGTATCTGGAGAGCCTGCGGGACGCGCGGGAGGTGTACCTCGACGGCAGCCGCGTCAAG
32881 -----+-----+-----+-----+-----+ 32940
CTCATAGACCTCTCGGACGCCCTGCGCGCCCTCCACATGGAGCTGCCGTGCGGCGAGTTC
4 E Y L E S L R D A R E V Y L D G S R V K -

GACGTCACCGCGCATCCCGCGTTCACAACCCGGCCCGGATGACGGCCCCGGCTGTACGAC
32941 -----+-----+-----+-----+-----+ 33000
CTGCAGTGGCGCGTAGGGCGCAAGGTGTTGGGCCGGGCTACTGCCGGGCCGACATGCTG
4 D V T A H P A F H N P A R M T A R L Y D -

AGCCTGCACGACCCCCGCCAGAAAGCGGTCTTGACGGCGCCACCGATGCCGGTGACGGT
33001 -----+-----+-----+-----+-----+ 33060
TCGGACGTGCTGGGGCGGGTCTTTCCGACGAGTCCCGCGGGTGCTACGGCCACTGCCA
4 S L H D P A Q K A V L T A P T D A G D G -

TTCACCCACCGCTTCTTACCGCACCGCGCAGCGTCGACGACCTGGTCAAGGACCAGGCC
33061 -----+-----+-----+-----+-----+ 33120
AAGTGGGTGGCGAAGAAGTGGCGTGGCGCGTGCAGCTGCTGGACCAGTTCCTGGTCCGG
4 F T H R F F T A P R S V D D L V K D Q A -

GCCATCGCATCCTGGGCGCGCAAGAGCTACGGCTGGATGGGGCGCAGCCCCGACTACAAG
33121 -----+-----+-----+-----+-----+ 33180
CGGTAGCGTAGGACCCGCGCGTTCGATGCCGACCTACCCCGCGTCGGGGCTGATGTTT
4 A I A S W A R K S Y G W M G R S P D Y K -

GCGTCGTTCTCGGCACGCTGGGGGCCAACGCCGACTTCTACGAGCCCTTCGCGGACAAC
33181 -----+-----+-----+-----+-----+ 33240
CGCAGCAAGGAGCCGTGCGACCCCGGTTGCGGCTGAAGATGCTCGGGAAGCGCCTGTTG
4 A S F L G T L G A N A D F Y E P F A D N -

GCCCGGCGCTGGTACCGGGAGTCGCAGGAGAAGGTGCTGTACTGGAACCATGCCTTCCTT
33241 -----+-----+-----+-----+-----+ 33300
CGGGCCGCGACCATGGCCCTCAGCGTCTCTTCCACGACATGACCTTGGTACGGAAGGAA
4 A R R W Y R E S Q E K V L Y W N H A F L -

CACCCGCCGGTTCGACCGCTCGCTGCCCGCCGACGAGGTGGGCGACGCTTTCATCCACGTC
33301 -----+-----+-----+-----+-----+ 33360
GTGGGCGGCCAGCTGGCGAGCGACGGGCGGCTGCTCCACCCGCTGCAGAAGTAGGTGCAG
4 H P P V D R S L P A D E V G D V F I H V -

GAGCGGGAGACCGACGCGGGCCTGGTGGTGAGCGGGGCCAAGGTCGTCGCGACCGGATCG
33361 -----+-----+-----+-----+-----+ 33420
CTCGCCCTCTGGGTGCGCCCGGACCACCACTCGCCCGGTTCCAGCAGCGCTGGCCTAGC
4 E R E T D A G L V V S G A K V V A T G S -

GCCCTCACCCACGCGCGTTTCATCTCGCACTGGGGACTTCCCATCAAGGACCGGAAGTTC
33421 -----+-----+-----+-----+-----+ 33480
CGGGAGTGGGTGCGCCGCAAGTAGAGCGTGACCCCTGAAGGGTAGTTCCTGGCCTTCAAG
4 A L T H A A F I S H W G L P I K D R K F -

GCCCTGGTGGCCACCGTGCCGATGGACGCGGACGGCCTCAAGGTGATCTGCCGTCCCTCC

22 D E I E R E V A Y A P G D A S A G R G R -
 CGGTCGTAGTTGACGACCGTGAAGTGCTCGGCGAGGAGACCGGCGAGCTTCTTGGCGTCG
 35101 -----+-----+-----+-----+-----+-----+-----+ 35160
 GCCAGCATCAACTGCTGGCACTTCACGAGCCGCTCCTCTGGCCGCTCGAAGAACCGCAGC
 22 R D Y N V V T F H E A L L G A L K K A D -
 GAGCGGTGCGCCAAGGCGGAGGCCACCAGGATCACCGCCGGCCCCCTCGCCCGACTTGTCG
 35161 -----+-----+-----+-----+-----+-----+-----+ 35220
 CTCGCCAGCCGGTTCGCGCTCCGGTGGTCCTAGTGGCGGCCGGGAGCGGGCTGAACAGC
 22 S R D A L A S A V L I V A P G E G S K D -
 AAGGCGATCGTGGTGCCGTGCGCCGATACCGTCGTTGATTCCACCTTGGCTGCTTTCTCA
 35221 -----+-----+-----+-----+-----+-----+-----+ 35280
 TTCCGCTAGCACCACGCGCAGCCGGCTATGGCAGCACTAAGGTGGAACCGACGAAAGAGT
 22 F I A T T G D A S V T T S E V K A A K E -
 CGGGTTGAAGACATAGCTTCCCTCAGATCACATTGTGGGGCGTGCTGCCGACAGTGGAGA
 35281 -----+-----+-----+-----+-----+-----+-----+ 35340
 GCCCAACTTCTGTATCGAAGGGAGTCTAGTGTAACACCCCGCACGACGGCTGTACCTCT
 22-< R T S S M -
 CCGGCGTCCGGAGGAAAAGTAATCGGTCTGCCAGAATTGGGGGTTCGGAGGGGCACGCC
 35341 -----+-----+-----+-----+-----+-----+-----+ 35400
 GGCCGCAGGCCTCCTTTTCATTAGCCAGGACGGTCTTAACCCCAAGGCCTCCCGTGCGG
 GACCGCTGCACGACGGCGCGCCCCGACCTTCCGGACATTGTCGTGCCCTCAGATGTGTTT
 35401 -----+-----+-----+-----+-----+-----+-----+ 35460
 CTGGCGACGTGCTGCCGCGGGGCTGGAAGGCCTGTAACAGCACGGGAGTCTACACAAA
 CGCATCTTCAGGAGTGCTCAGTGATCCGTGAGGTGAGAAAGGGACGGTGGTCCGGTCAGT
 35461 -----+-----+-----+-----+-----+-----+-----+ 35520
 GCGTAGAAGTCTCTACGAGTCACTAGGCACTCCACTCTTCCCTGCCACCAGGCCAGTCA
 18-* * -
 CGTTGCCGCGCGGGCTGTTCTGGTAAGCGGCCAGACGCCACTGCCCGTCTCTGTTTCGACGG
 35521 -----+-----+-----+-----+-----+-----+-----+ 35580
 GCAACGGCGCGCCCCGACAAGACCATTTCGCCGGTCTGCGGTGACGGGCAGGACAAGCTGCC
 18 D N G R P S N Q Y A A L R W Q G D Q E V -
 CCAGCCAGGAGGCCCGGACGGCGCCGTCGCCGCTCGCCTCGGTCTCCCCCGGGGCGAGGA
 35581 -----+-----+-----+-----+-----+-----+-----+ 35640
 GGTCCGTCTCTCCGGGCTGCCGCGGCAGCGGCGAGCGGAGCCAGAGGGGGCCCCGCTCCT
 18 A L W S A R V A G D G S A E T E G P A L -
 TGCCGCCCTCGGTGATGAGCAGGGCGATGCCGTCGCCGAGCAGGCGCGCTCGATGGGGC
 35641 -----+-----+-----+-----+-----+-----+-----+ 35700
 ACGGCGGGAGCCACTACTCGTCCCGCTACGGCAGCGGCTCGTCCGCGCGCAGCTACCCCG
 18 I G G E T I L L A I G D G L L R A D I P -
 TGCCGATGACACGGGTGCCCTTGTACGGGCCCCGGAAGGCGGCCGCCATGTGGGTGCGGA
 35701 -----+-----+-----+-----+-----+-----+-----+ 35760
 ACGGCTACTGTGCCCACGGGAACATGCCCGGGCGCTTCCGCGGCGGTACACCCACGCCT
 18 S G I V R T G K Y P G A F A A A M H T R -
 TGTTCTCGCGGCCCTTGCGGAAGAGGCCGGGGAGGATCATCGTCCCGTCTCTGGCGAAGA
 35761 -----+-----+-----+-----+-----+-----+-----+ 35820
 ACAAGAGCGCCGGAACGCCTTCTCCGGCCCCCTCTAGTAGCAGGGCAGGAGCCGCTTCT
 18 I N E R G K R F L G P L I M T G D E A F -
 CGTCGGCGAACCAGGTGCGGCTCGTGGTTCGGCCCAGGCGGCCACGATGCGCGCCGGCAGAG
 35821 -----+-----+-----+-----+-----+-----+-----+ 35880
 GCAGCCGCTTGGCCAGCCGACGACAGCCGGGTCCGCCGGTGCTACGCGCGGGCGCTCTC
 18 V D A F R D A D H D A W A A V I R A P L -
 CGGCTACCGCTGCCAGGGCGGCGTCCGGAGCGGAGGTGGTTCGAGTCGGTGCTGGTCATAT

35881 -----+-----+-----+-----+-----+ 35940
 18-< GCCGATGGCGACGGTCCCGCCGAGCCCTCGCCTCCACCAGCTCAGCCACGACCAGTATA
 A A V A A L A A D P A S T T S D T S T M -
 CGCGGTTCCCGTCCGTTGGTTGGCGGTTTCGGCACGGCCCGCAGCCCTGCCCCAGCCCCGA
 35941 -----+-----+-----+-----+-----+ 36000
 GCGCCAAGGGCAGGCAACCAACCGCCAAAGCCGTGCCGGGCGTCGGGACGGGCTCGGGCT
 CGCTGGCAGGCGGCCCCGTGTCATCAGGCATCTCCTGCGTTGCGCCCCACGCCAGTCACTTC
 36001 -----+-----+-----+-----+-----+ 36060
 GCGACCGTCCGCGGGGGCAGTAGTCCGTAGAGGACGCAACGCGGGGTGCGGTCACTGAAG
 ACGGCCAGAACAAAGTCGCGCATTCTGGAAGAAGCTGAGGCCCGCGACCCGGTGGCAGCAT
 36061 -----+-----+-----+-----+-----+ 36120
 TGCCGGTCTTGTTTCAGCGCGTAAGACCTTCTTCGACTCCGGGCGCTGGGCCACGCTGCTA
 CTGCGGTGTACCGAGTTTCGCACACGTTTACGCACGGAGGCTCGATGCCCCGCTGTCAATG
 36121 -----+-----+-----+-----+-----+ 36180
 5-> GACGCCACAGTGCCTCAAGCGTGTGCAAATGCGTGCCTCCGAGCTACGGGCGACAGTTAC
 M P A V N G -
 GATCGGTGCAGTCAGGCCAGTCGCACCGACGCTCCGTCGTGGCGACGGTGGTGGGCAACT
 36181 -----+-----+-----+-----+-----+ 36240
 5 CTAGCCACGTCACTCCGGTCAGCGTGGCTGCGAGGCAGCACCGCTGCCACCACCCGTTGA
 S V Q S G Q S H R R S V V A T V V G N F -
 TCGTGGAGTCGTTGACTGGCTCGCCTACGGGCTCTTCGCTCCTCTCTTCGCGGCTCAGT
 36241 -----+-----+-----+-----+-----+ 36300
 5 AGCACCTCAGCAAGCTGACCGAGCGGATGCCCCGAGAAGCGAGGAGAGAAGCGCCGAGTCA
 V E S F D W L A Y G L F A P L F A A Q F -
 TCTTCCCCCTCGTCCAACAGTTACCTCCCTGCTCGGCGCGTTCGCGGTCTTCGGCACGG
 36301 -----+-----+-----+-----+-----+ 36360
 5 AGAAGGGGAGCAGGTTGGTCAAGTGGAGGGACGAGCCGCGCAAGCGCCAGAAGCCGTGCC
 F P S S N Q F T S L L G A F A V F G T G -
 GCATGCTCTTCCGGCCGATCGGCGGGGTCTGCTGGGCGGCCCTCGCCGACCGGCGCGGCC
 36361 -----+-----+-----+-----+-----+ 36420
 5 CGTACGAGAAGGCCGGCTAGCCGCCCCAGGACGACCCGGCGGAGCGGCTGGCCGCGCCGG
 M L F R P I G G V L L G R L A D R R G R -
 GCGCCCCCGCCCTGATGCTGGCGATCGGACTGATGACCGGCGGCTCGACCCTGATCGCCG
 36421 -----+-----+-----+-----+-----+ 36480
 5 CCGCGGGGCGGGACTACGACCGCTAGCCTGACTACTGGCCGCGGAGCTGGGACTAGCGGC
 R P A L M L A I G L M T G G S T L I A V -
 TCGTCCCCACCTACGAGCACATCGGGATCCTCGCCCCGCTGCTTCTGCTGCTCGCCCCGC
 36481 -----+-----+-----+-----+-----+ 36540
 5 AGCAGGGGTGGATGCTCGTGTAGCCCTAGGAGCGGGGCGACGAAGACGACGAGCGGGCCG
 V P T Y E H I G I L A P L L L L L A R L -
 TCGCCAGGGAGTCTCCTCGGGCGGGGAATGGACAGCGGCGGCCACCTACCTGATGGAGA
 36541 -----+-----+-----+-----+-----+ 36600
 5 AGCGGGTCCCTCAGAGGAGCCCCGCCCTTACCTGTGCGCCGCGGTGGATGGACTACCTCT
 A Q G V S S G G E W T A A A T Y L M E I -
 TCGCGCCGAAGAACC GCCGGTGCCTCTACAGCAGCCTCTTCTCCGTGACGACCATGGCGG
 36601 -----+-----+-----+-----+-----+ 36660
 5 AGCGCGGCTTCTTGGCGGCCACGGAGATGTGCTCGGAGAAGAGGCACTGCTGGTACCGCC
 A P K N R R C L Y S S L F S V T T M A G -
 GCCCCTTCGTGCGATCGCTGCTGGGCGCGGGCCTCGGCGTGTGGCTGGGAACCGCGACGA
 36661 -----+-----+-----+-----+-----+ 36720
 5 CGGGGAAGCAGCGTAGCGACGACCCGCGCCCGGAGCCGCACACCGACCCTTGGCGTGTCT
 P F V A S L L G A G L G V W L G T A T M -

```

      TGGAGGCCTGGGGCTGGCGGGTGCCGTTCTCCTCGGCGGCGTCTTCGGCGTGATCCTGC
36721 -----+-----+-----+-----+-----+-----+ 36780
      ACCTCCGGACCCCGACCGCCACGGCAAGGAGGAGCCGCCGAGAAGCCGCACTAGGACG
5      E A W G W R V P F L L G G V F G V I L L -

      TGTTCTCGCGCGTCTGGCTCACCGAGACCGAGGTCTTCGCCCGGGAGGTGCGGCCCCGGG
36781 -----+-----+-----+-----+-----+-----+ 36840
      ACAAGGACGCGGCAGCCGAGTGGCTCTGGCTCCAGAAGGCGGCCCTCCACGCCGGGGCCCC
5      F L R R R L T E T E V F R R E V R P R A -

      CCCGGCGCGGCTCACTGGGCCAGCTGATCGGAGCCACCGCCCCAGGTGCTGCTGGCCG
36841 -----+-----+-----+-----+-----+-----+ 36900
      GGGCCGCGCCGAGTGACCCGGTTCGACTAGCCTCGGGTGGCGGGGTCCACGACGACCGGC
5      R R G S L G Q L I G A H R P Q V L L A V -

      TGATGCTGGTGGCCGGAAGTGGGCGTCATCGGCGGAACGTGGTTCGACCGCGGTCCCGGCGA
36901 -----+-----+-----+-----+-----+-----+ 36960
      ACTACGACACCGGCCTGACCCGAGTAGCCGCCTTGACACAGCTGGCGCCAGGGCCGCT
5      M L V A G L G V I G G T W S T A V P A M -

      TGGGCCACCGTCTGATCGGCTCGCAGACGATGTTCTGGGTGGTGGTCTGTGTGACCGGCT
36961 -----+-----+-----+-----+-----+-----+ 37020
      ACCCGGTGGCAGACTAGCCGAGCGTCTGCTACAAGACCCACACAGACACACTGGCCGA
5      G H R L I G S Q T M F W V V V C V T G S -

      CGGTTCATCCTGCTGCAGGTACCCATAGGGCTGCTCGCCGACCGGGTGAACCGGGCAGGT
37021 -----+-----+-----+-----+-----+-----+ 37080
      GCCAGTAGGACGACGTCCATGGGTATCCCGACGAGCGGCTGGCCACCTTGCCCCGTCCA
5      V I L L Q V P I G L L A D R V E P G R F -

      TCCTGATCGTCTCCAGCGTCGTCTTCGCCGCTGTGGGCTCGTACGCCTACCTACCGTCC
37081 -----+-----+-----+-----+-----+-----+ 37140
      AGGACTAGCAGAGGTTCGAGCAGAAGCGGCGACACCCGAGCATGCGGATGGAGTGGCAGG
5      L I V S S V V F A A V G S Y A Y L T V Q -

      AGGACTCCTTCGCGAGCCTGGCGTTTACGTACAGCACCGGAGTGATCTTCCTCGGCTGCG
37141 -----+-----+-----+-----+-----+-----+ 37200
      TCCTGAGGAAGCGCTCGGACCGCAAGTGCATGTCTGTCGCTCACTAGAAGGAGCCGACGC
5      D S F A S L A F T Y S T G V I F L G C V -

      TCACCATGGTGTGCTGCCGAAGATGCTCTCCAGAATCTTCCTCCGAGATACGCGGCCTGG
37201 -----+-----+-----+-----+-----+-----+ 37260
      AGTGGTACCACGACGGCTTCTACGAGAGGTCTTAGAAGGGAGGCGTCTATGCGCCGGACC
5      T M V L P K M L S R I F P P Q I R G L G -

      GCATCGGGCTGCCGCACGCCTCGACCACCGCACTCCTCGGCGGGGCGGGGCCACTGCTGG
37261 -----+-----+-----+-----+-----+-----+ 37320
      CGTAGCCCGACGGCGTGGGAGCTGGTGGCGTGAGGAGCCGCCCGCCCGGTGACGACC
5      I G L P H A S T T A L L G G A G P L L A -

      CCGCCTACTCCGACGAGCGAGGCGCCTCGGGCTGGTTCATCGCCGCCGTGATGGCCGCGG
37321 -----+-----+-----+-----+-----+-----+ 37380
      GCGGATGAGGCTGCTCGCTCCGCGGAGCCCGACCAAGTAGCGGCGGCACTACCGGCGCC
5      A Y S D E R G A S G W F I A A V M A A V -

      TCCTGCTCGCCTGGCCGGCCACCCTGTGGGAGCGACGGCTGTTCGCGCCCGGACGGCCC
37381 -----+-----+-----+-----+-----+-----+ 37440
      AGGACGAGCGGACCGGCCGGTGGGACACCCTCGCTGCCGACAAGGCGCGGGCCTGCCGGG
5      L L A W P A T L W E R R L F R A R T A P -

      CGGGAAGCGAGCCGGTTCCCGAATCCGCCGTCGCCCCCCCCGTCGGGTGACCGTCCGCAC
37441 -----+-----+-----+-----+-----+-----+ 37500
      GCCCTTCGCTCGGCCAAGGGCTTAGGCGGCAGCGGGCGGGGCGAGCCCACTGGCAGGCGTG
5-*      G S E P V P E S A V A R P V G * -

```

37501 TTCTGCATCCCGTCCGGCACCGAGCGCCGGCGACCTTCCCGACTGAGAGGTTGACATCAT
 -----+-----+-----+-----+-----+ 37560
 AAGACGTAGGGCAGGCCGTGGCTCGCGGCCGCTGGAAGGGCTGACTCTCCAACCTGTAGTA
 23-> M -

 37561 GACGACGTCCGACACCACCGACCGGTCCCAGGACGGCGTGCCGCCGCTCTCCTTCCACCA
 -----+-----+-----+-----+-----+ 37620
 CTGCTGCAGGCTGTGGTGGCTGGCCAGGGTCCTGCCGCACGGCGGCGAGAGGAAGGTGGT
 23 T T S D T T D R S Q D G V P P L S F H Q -

 37621 GGAGTTCCTGTGCATGTTTCGACAGCGGGAACGACGGCGCCGACGTGGGGCCGTTTCGGCCCC
 -----+-----+-----+-----+-----+ 37680
 CCTCAAGGACACGTACAAGCTGTGCGCCCTTGCTGCCGCGGCTGCACCCCGCAAGCCGGG
 23 E F L C M F D S G N D G A D V G P F G P -

 37681 CATGTACCACATCGTCGGAGCCTGGCGGCTGACCGGCGGGATCGACGAGGAGACCTTGCG
 -----+-----+-----+-----+-----+ 37740
 GTACATGGTGTAGCAGCCTCGGACCGCCGACTGGCCGCCCTAGCTGCTCCTCTGGGACGC
 23 M Y H I V G A W R L T G G I D E E T L R -

 37741 CGAGGCGCTGGGTGACGTCGTCGTGCGCCACGAGGCCCTGCGCACATCGTGGTCCGCGA
 -----+-----+-----+-----+-----+ 37800
 GCTCCGCGACCCACTGCAGCAGCACGCGGTGCTCCGGGACGCGTGTAGCGACCAGGCGCT
 23 E A L G D V V V R H E A L R T S L V R E -

 37801 AGGTGGCACGCACCGGCCGGAGATCCTGCCTGCGGGGCCCGCCGCGCTGGAGGTCCGTGA
 -----+-----+-----+-----+-----+ 37860
 TCCACCGTGCCTGGCCGGCCCTCTAGGACGCGACGCCCCGGGCGGCGACCTCCAGGCACT
 23 G G T H R P E I L P A G P A A L E V R D -

 37861 TCTCGGCGACGTCGACGAGTCGGAGCGGGTGCGGCGCGGTGAGGAACTGCTCAACGAGGT
 -----+-----+-----+-----+-----+ 37920
 AGAGCCGCTGCAGCTGCTCAGCCTCGCCCCACGCCGCGCCACTCCTTGACGAGTTGCTCCA
 23 L G D V D E S E R V R R G E E L L N E V -

 37921 GGAGTCGACCGGTCTGAGCGTGCGGGAGCTGCCCCCTGCTGCGGGCCGTGCTCGGACGCTT
 -----+-----+-----+-----+-----+ 37980
 CCTCAGCTGGCCAGACTCGCACGCCCTCGACGGGGACGACGCCCGGCACGAGCTGCGAA
 23 E S T G L S V R E L P L L R A V L G R F -

 37981 CGACCAGAAGGACGCGGTGCTGGTCTCATCGCCACACACCGCCGCGGACGCTGGGC
 -----+-----+-----+-----+-----+ 38040
 GCTGGTCTTCTGCGCCACGACCAGGAGTAGCGGGTGGTGTGGCGGCGCTGCGGACCCG
 23 D Q K D A V L V L I A H H T A A D A W A -

 38041 CATGCACGTCATCGCCCGGACCTGCTCAACCTGTACGCCGCCAGGCGCGGGAACCCGGT
 -----+-----+-----+-----+-----+ 38100
 GTACGTGCAGTAGCGGGCGCTGGACGAGTTGGACATGCGGCGGTCCGCGCCCTTGGGCCA
 23 M H V I A R D L L N L Y A A R R G N P V -

 38101 TCCCCCGCTCCCCGAGCCGGGCCAGCATGCCGAGTTCGCCCGCTGGGAGCGCGAGGCGGC
 -----+-----+-----+-----+-----+ 38160
 AGGGGGCGAGGGGCTCGGCCGGGTCTGACGGCTCAAGCGGGCGACCCTCGCGCTCCGCCG
 23 P P L P E P A Q H A E F A R W E R E A A -

 38161 CGAGGCACCGCGGGTCGCGGTCTCGAAGGAATTCTGGCGCAAGCGCTCCAGGGCGCGCG
 -----+-----+-----+-----+-----+ 38220
 GCTCCGTGGCGCCAGCGCCAGAGCTTCTCTTAAGACCGCGTTTCGCGGAGGTCCCGCGCGC
 23 E A P R V A V S K E F W R K R L Q G A R -

 38221 GATCATCGGGCTGGAGACGGACATACCGCGCTCGGCGGGGCTGCCCAAGGGCACCGCGTG
 -----+-----+-----+-----+-----+ 38280
 CTAGTAGCCCGACCTCTGCCTGTATGGCGCGAGCCCGCCCGACGGTTCCCGTGGCGCAC
 23 I I G L E T D I P R S A G L P K G T A W -

 GCAGCGCTTCGCCGTACGCGGGAACTGGCCGACGCCGTGGTGGAGTTCTCACGGGCCGC

26 TGCAGGCCACTCCTTCCGCGCTGCAAGTAGCTCGAGCCTCCCGTCAGCCAGAGCCGGCAC
T S G E E G A T F I E L G G Q S V S A V -

39121 CGCATCGCCACGCGTATCCAGGAGGAGCTCGACATCTGGGTCGACATCGGCGTCTCTTTC
-----+-----+-----+-----+-----+ 39180
26 GCGTAGCGGTGCGCATAGGTCCTCCTCGAGCTGTAGACCCAGCTGTAGCCGAGGAGAAG
R I A T R I Q E E L D I W V D I G V L F -

39181 GACGACCCGGATCTGCCTACCTTCATCGCGGCGGTCTCGGACGGCCGACGCCGCGGGC
-----+-----+-----+-----+-----+ 39240
26 CTGCTGGGCCTAGACGGATGGAAGTAGCGCCGCCAGCAGGCCTGCCGGCTGCGGCGCCCG
D D P D L P T F I A A V V R T A D A A G -

39241 GGCGAGGGCTCCGGAACGCAGTGAGACTCGCCGGGCGCCGTCTCCCCGCGGCGCCCGGTT
-----+-----+-----+-----+-----+ 39300
26-* CCGCTCCCGAGGCCTTGCCTCACTCTGAGCGGCCCGCGGCAGAGGGGCGCCGCGGGCCAA
G E G S G T Q * -

39301 TCACATGGCTGAGGCGGTTACCCGGTACCGGTTGAACCGCCTCAGCCATGTGAAACCGG
-----+-----+-----+-----+-----+ 39360
AGTGTACCGACTCCGCCAAGTGGGCCATGGCCCACTTGGCGGAGTCGGTACACTTTGGCC

39361 GCCTGGTCAGCGCAGCTGGATGTCCGTCTCCCGGCGATCGCCCGAGGAACTCGCCGCG
-----+-----+-----+-----+-----+ 39420
24-* CGGACCAGTCGCGTCGACCTACAGGCAGAGGGCCCGCTAGCGGGCCTCCTTGAGCGGCGC
* R L Q I D T E R A I A R L F E G R -

39421 GGACAGCGCGTCCGGCGACCAGCTCGATGTCTCGGCCATGTACCGGTGACGCCCAGCGT
-----+-----+-----+-----+-----+ 39480
24 CCTGTGCGCGCAGCCGCTGGTTCGAGCTACAGCAGCCGGTACATGGCCAGCTGCGGGTTCGCA
S L A D A V L E I D D A M Y R D V G L T -

39481 CGGAACCAGCCGGCGCACCGCTTCGTACGTGGCCTTCGCCGCCGGGCTCAAGCCGTCGAA
-----+-----+-----+-----+-----+ 39540
24 GCCTTGGTTCGGCCGCGTGGCGAAGCATGCACCGGAAGCGGCGCCGAGTTTCGGCAGCTT
P V L R R V A E Y T A K A A P S L G D F -

39541 CCGGCCGGAGATGTTCGACCGCCTGGGCGGCGGCCAGGTACTCCACCGCAGGATCTTGTT
-----+-----+-----+-----+-----+ 39600
24 GGCCGGCCTCTACAGCTGGCGGACCCGCGCCGGTCCATGAGGTGGCGCTCCTAGAACAA
R G S I D V A Q A A A L Y E V A L I K N -

39601 GTTGTTCGACAGGACCCGGCGGGCGTTGCGGGCCGAGATCAGGCCCATGCTCACCACGTC
-----+-----+-----+-----+-----+ 39660
24 CAACAAGCTGTCTGGGCGCCCGCAACGCCCGGCTCTAGTCCGGGTACGAGTGGTGCAG
N N S L V R R A N R A S I L G M S V V D -

39661 CTGGTTGTTCGCGTGGACGGGACGCTCTGGGTGCTGGCCGGGCGGATCGTCCGGTTCTC
-----+-----+-----+-----+-----+ 39720
24 GACCAACAGCGGCAACCTGCCCTGCGAGACCCACGACCGGCCCGGCTAGCAGGCCAAGAG
Q N D G N S P V S Q T S A P G I T R N E -

39721 GGCCACCAGTGCAGTGGCCGGGTTACTGGGCGCCGGCGAATCCGCTGTGCAGCCCCGGGTC
-----+-----+-----+-----+-----+ 39780
24 CCGGTGGTCACGCCACCGGCCCATGACCCGCGGCGGCTTAGGCGACACGTCGGGGCCCAG
A V L A T A P Y Q A G A F G S H L G P D -

39781 CCCGGAGACGAGGAACTCCGGGAGGCCGTAGCTGAGGTGCCGGTTCAGGACCCGGTTGAT
-----+-----+-----+-----+-----+ 39840
24 GGGCCTCTGCTCCTTGAGGCCCTCCGGCATCGACTCCACGGCCAAGTCTGGGCCAACTA
G S V L F E P L G Y S L H R N L V R N I -

39841 CTGCCGCTCGGCCAGGACGCCGAGCTGGGTGAGCGCGATGGTCACGAAGTCCATCGCGAA
-----+-----+-----+-----+-----+ 39900
GACGGCGAGCCGGTCTGCGGCTCGACCCACTCGCGCTACCAGTGCTTACAGGTAGCGCTT

24 Q R E A L V G L Q T L A I T V F D M A F -
CGCGATCGGCTGACCGTGGAAGTTTCGCCCCGTGGAAGATCTCCTTGCCCTCGAAGAAGAG
39901 -----+-----+-----+-----+-----+-----+ 39960
GCGCTAGCCGACTGGCACCTTCAAGCGGGGCACCTTCTAGAGGAACGGGAGCTTCTTCTC
24 A I P Q G H F N A G H F I E K G E F F L -
CGGGTTGTCTGGCCGAGTTGAGCTCGATGCGCAGCTTGTGCCGCGCGTGGTACAAGGT
39961 -----+-----+-----+-----+-----+-----+ 40020
GCCCAACAGCAACCGGCTCAACTCGAGCTACGCGTCGAACACGGCGCGCACCATGTTCCA
24 P N D N A S N L E I R L K H R A H Y L T -
GTCGCGCACCGCCCCGACGACCTGGGGGATGGCCCGCAGCGAGTAGGCCTTCTGCAGGTA
40021 -----+-----+-----+-----+-----+-----+ 40080
CAGCGCGTGGCGGGGCTGCTGGACCCCTACCGGGCGTCGCTCATCCGAAGACGTCCAT
24 D R V A G V V Q P I A R L S Y A K Q L Y -
GATCTCCGAGCGCTGGACGTCCTTGCCGGCCTCCTTGTCTTCTGGAGTTCTCGGCGCAG
40081 -----+-----+-----+-----+-----+-----+ 40140
CTAGAGGCTCGCGACCTGCAGGAACGGCCGGAGGAACAGGAAGACCTCAAGAGCCGCGTC
24 I E S R Q V D K G A E K D K Q L E R R L -
GTCGGCGTGTCTCGACCGTCAGTCCGCTGCCCCGCATCAGGGCCCCGCATGTTGGCGGCGGT
40141 -----+-----+-----+-----+-----+-----+ 40200
CAGCCGCACGAGCTGGCAGTCAGGCGACGGGGCGTAGTCCCGGGCGTACAACCGCCGCCA
24 D A H E V T L G S G R M L A R M N A A T -
GTCGATCTGGCCCTCGTGCGGGCGGGCTATGTCTGCCCCCTCCGCGAGGAAGGGGCTGGT
40201 -----+-----+-----+-----+-----+-----+ 40260
CAGCTAGACCGGGAGCACGCCCCGCGGATACAGCACGGGGAGGCGCTCCTTCCCCGACCA
24 D I Q G E H P R A I D H G E A L F P S T -
CGATCCGCGTACCGCCTCGATGAGCAGAGCCGTCACGATCTCGGCCTGCTGGGCCTGCTC
40261 -----+-----+-----+-----+-----+-----+ 40320
GCTAGGCGCATGGCGGAGCTACTCGTCTCGGCAGTGCTAGAGCCGGACGACCCGGACGAG
24 S G R V A E I L L A T V I E A Q Q A Q E -
CAGGGCCCGTCCGACGACCAGGGAGCCAGACCGGTATCCCGGACGTGCCGTTGATCAG
40321 -----+-----+-----+-----+-----+-----+ 40380
GTCCCCGGGACAGGCTGCTGGTCCCTCGGGTCTGGCCAGTAGGGCCTGCACGGCAACTAGTC
24 L A R G V V L S G L G T M G S T G N I L -
TGCGAGGCCCTCCTTGAAGCGCAGTTTCGAGCGGCTCGATGCCCCGCTCGGCCAGCACCTG
40381 -----+-----+-----+-----+-----+-----+ 40440
ACGCTCCGGGAGGAACTTCGCGTCAAGCTCGCCGAGCTACGGGGCGAGCCGGTCTGGAC
24 A L G E K F R L E L P E I G R E A L V Q -
GGCGGTCTCCACCGGCCGTCCGTGCGCAGGACGTAGCCCTCTCCGATGAGGGTGCTCGC
40441 -----+-----+-----+-----+-----+-----+ 40500
CCGCCAGAGGTGGCCGGCAGGCAGCGCTCCTGCATCGGGAGAGGCTACTCCACGAGCG
24 A T E V P R G D R L V Y G E G I L T S A -
GACGTGGGAGAGGGGAGCCAGGTGCGCGCTCGCCCCGAGTGACCCGATCTCGGGTATGGC
40501 -----+-----+-----+-----+-----+-----+ 40560
CTGCACCCTCTCCCCCTCGGTCCAGCGGCGAGCGGGCTCACTGGGCTAGAGCCCATAACCG
24 V H S L P A L D G S A G L S G I E P I A -
CGGGGTGATGCCCTCGTTTCAAGTACTGCGCGAGGCGTTTCGAGGATGATGGGGCGCACCGC
40561 -----+-----+-----+-----+-----+-----+ 40620
GCCCCACTACGGGAGCAAGTCCATGACGCGCTCCGCAAGCTCCTACTACCCCGCGTGGCG
24 P T I G E N L Y Q A L R E L I I P R V A -
GGAGTGGCCCTTGGCGAGGGTGTTTCAAGCGGGCGGCGACGATCGCCCGCGCCTCGTCTC
40621 -----+-----+-----+-----+-----+-----+ 40680
CCTCACCGGGAACCGCTCCCAAGTCGGCCCGCGCTGCTAGCGGGCGCGGAGCAGGAG

005070" 0070460

24 S H G K A L T N L R A A V I A R A E D E -
GGCGAACAGCGGACCGACTCCCGCGCTGTGGCTACGGACGAGATTGGTCTGCAGTTCGAC
40681 -----+-----+-----+-----+-----+ 40740
CCGCTTGTGCGCTGGCTGAGGGCGCGACACCGATGCCTGCTCTAACCAGACGTCAAGCTG
24 A F L P G V G A S H S R V L N T Q L E V -
TTCCTTCGACTTGTGCGACCTGCATGTAGATCATCTCGCCGTACCCGGTGGTCACCCCGTA
40741 -----+-----+-----+-----+-----+ 40800
AAGGAAGCTGAACAGCTGGACGTACATCTAGTAGAGCGGCATGGGGCCACAGTGGGGCAT
24 E K S K D V Q M Y I M E G Y G T T V G Y -
GATGGGGATGTTCTGTTGCGCGATCCCTTCGAAGATCTCCCGGCTCTTCTGGGCGCTTCGC
40801 -----+-----+-----+-----+-----+ 40860
CTACCCCTACAAGACAAGCCGCTAGGGAAGCTTCTAGAGGGCCGAGAAGACCCGGAAGCG
24 I P I N Q E A I G E F I E R S K Q A K A -
GATGGATTTCGGCCGGTACGTCGACCGTCGCGCGTTCCTCCGCGACGCGGCGTACGGCTTC
40861 -----+-----+-----+-----+-----+ 40920
CTACCTAAGCCGGCCATGCAGCTGGCAGCGCGCAAGGAGGCGCTGCGCCGCATGCCGAAG
24 I S E A P V D V T A R E E A V R R V A E -
GACGGTCAGGGTCTCGCCGTCGACGGAACCCGGGACGATCTCGGTCTCGACTTGAGTCAA
40921 -----+-----+-----+-----+-----+ 40980
CTGCCAGTCCCAGAGCGGCAGCTGCCTTTGGCCCTGCTAGAGCCAGAGCTGAACTCAGTT
24 V T L T E G D V S V P V I E T E V Q T L -
TGCCATCACTCCATGGGTAGCGGCCGAGGCCGGTGTACGACAGGTCAGGGGGTGGGTTCG
40981 -----+-----+-----+-----+-----+ 41040
ACGGTAGTGAGGTACCCATCGCCGGCTCCGGCCACATGCTGTCCAGTCCCCCACCCAAGC
24-< A M -
TGAGGCGCGGCTCAGCGGGTGAGCCGGGAGCGGTCCACCTTCCCCGCGGCGTTGCGCGGC
41041 -----+-----+-----+-----+-----+ 41100
ACTCCGCGCCGAGTCGCCCCTCGGCCCTCGCCAGGTGGAAGGGGCGCCGAACGCGCCG
25-* * R T L R S R D V K G A A N R P -
AGGCGTGAAGTCAGGCGGGTGAAGACGGCGGGCAGTGCGAGGGGGCCGAACCTGGCCGCGC
41101 -----+-----+-----+-----+-----+ 41160
TCCGCACTTCAGTCCGCCCACTTCTGCGCCCGTCACGCTCCCCGGCTTGACCGGCGCG
25 L R S T L R T F V A P L A L P G F Q G R -
AGATGGGAACGCCAGGCCCGGATGTCCGCGCGCACGTCTCTCCGGCCCTCTCCTTGTGGC
41161 -----+-----+-----+-----+-----+ 41220
TCTACCCTTGGGTCCGGGCTACAGGCGCGCGTGCAGGAGGGCCGGAGAGGAACACCG
25 L H S R W A R I D A R V D E R G E G Q P -
ACCACGTACACGGCGAGGCGGGTCAACAGGCCCTGGCCGTTGACGTGGGGGAGGACCGCG
41221 -----+-----+-----+-----+-----+ 41280
TGGTGCATGTGCCGCTCCGCCCAGTGGTCCGGGACCGGCAACTGCACCCCTCTTGGCGC
25 V V Y V A L R T V L G Q G N V H P L V A -
CACTCCAGGACCGAGGGTCAAGGTTTTCAGCGCGGCTCGATCTCGGTGAGTTCCAAGCGG
41281 -----+-----+-----+-----+-----+ 41340
GTGAGGTCCTGGCTCCCCAGTGCCAAGTCGCGCCGAGCTAGAGCCACTCAAGGTTCCGC
25 C E L V S P D R N L A A E I E T L E L R -
TTCCCGAACAGCTTGACCTGGAAGTCTTGGCGCCCCGGAATTCCAGGGCTCCGTGGAAC
41341 -----+-----+-----+-----+-----+ 41400
AAGGGCTTGTGCAACTGGACCTTCAGGAACGCCGGGGCTTAAGGTCCCAGGCAGCTTG
25 N G F L K V Q F D K R G R F E L A G D F -
CGTACCCGCGCCAGATCCCCGGTCCGGTACACCGGTACCGTCCGGGGCGAGGCCGGCG
41401 -----+-----+-----+-----+-----+ 41460
GCATGGGCGCGGTCTAGGGGCCAGGCCATGGTGGCCAGTGGCAGGCCCGCTCCGGCCGC
25 R V R A L D G T R Y W R D G D P A L G A -

[illegible]

	CGTGGAAAGTTGGCTCCTTGCCGCTGGCTCATCATCGCGATGGCGCTGGGCCCGCTG	
43960	+-----+-----+-----+-----+	44019
	GCACCTTTCACCGACCAGGACCGGGCGGACCGAGTAGTAGCGCTACCGCGACCCGGGCGAC	
27	R G K W L V L A A W L I I A M A L G P L -	
	GCGGGGAAGCTCGCCGACTCCAGGACTCCAGCGCCAACGCCTTCCTTCCGCGCAGCTCG	
44020	+-----+-----+-----+-----+	44079
	CGCCCCTTCGAGCGGTGCAAGTCTGAGGTTCGCGGTTGCGGAAGAAGGCGCGTCGAGC	
27	A G K L A D V Q D S S A N A F L P R S S -	
	GAGTCCGCGAAGCTGAACAAGGAAGTGGAGAAGTTCGCGCCGACGAGCTGATGCCGGCC	
44080	+-----+-----+-----+-----+	44139
	CTCAGGCGCTTCGACTTGTTCTTGCCTCTTCAAGGCGCGGCTGCTCGACTACGGCCGG	
27	E S A K L N K E L E K F R A D E L M P A -	
	GTGGTGGTCTACAGCGCCGACGGCTCGCTGCCCGCCGAGGGGCGGGCCAAGGCCGAGAAG	
44140	+-----+-----+-----+-----+	44199
	CACCACCAGATGTCGCGGCTGCCGAGCGACGGGCGGCTCCCCGCCCGGTTCCGGCTCTTC	
27	V V V Y S A D G S L P A E G R A K A E K -	
	GACATAGCCGCTTCCAGGAGCTGGCCGCCGAGGGCGAGAAGGTGGAAGCGCCCTGGAG	
44200	+-----+-----+-----+-----+	44259
	CTGTATCGGCGGAAGGTCTTCGACCGGCGGCTCCCGCTCTTCCAGCTTCGCGGGGACCTC	
27	D I A A F Q E L A A E G E K V E A P L E -	
	TCCGAGGACGGCCAGGCGCTCATGGTCGTCGTTCCGCTGATCAGCGACGCCGACATCGTC	
44260	+-----+-----+-----+-----+	44319
	AGCCTCCTGCCGGTCCGCGAGTACCAGCAGCAAGGCGACTAGTCGCTGCGGCTGTAGCAG	
27	S E D G Q A L M V V V P L I S D A D I V -	
	GCCACGACGAAGAAGGTCCGCGATGTCGCGGACGCCAACGCCCCCCGGGCGTCGCCATC	
44320	+-----+-----+-----+-----+	44379
	CGGTGCTGCTTCTTCCAGGCGCTACAGCGCCTGCGGTTGCGGGGGGGCCCGCAGCGGTAG	
27	A T T K K V R D V A D A N A P P G V A I -	
	GAGGTGGGCGGGCCCGCCGGTTCGACGACCGACGCCGCCGCGCTTTCGAGTCCCTCGAC	
44380	+-----+-----+-----+-----+	44439
	CTCACCCGCCCCGGGCGGCCAGCTGCTGGCTGCGGCGGCCGGAAGCTCAGGGAGCTG	
27	E V G G P A G S T T D A A G A F E S L D -	
	TCCATGCTGATGATGGTCACCGGCCTTGTGGTTCGCCATCCTGCTGCTGATCACCTACCGC	
44440	+-----+-----+-----+-----+	44499
	AGGTACGACTACTACAGTGGCCGGAACACCAGCGGTAGGACGACGACTAGTGGATGGCG	
27	S M L M M V T G L V V A I L L L I T Y R -	
	TCCCCATCCTGTGGCTGCTGCCCCCTGCTCTCCGTTCGGCTTCGCCTCCGTGCTGACCCAG	
44500	+-----+-----+-----+-----+	44559
	AGGGGGTAGGACACCGACGACGGGGACGAGAGGCAGCCGAAGCGGAGGCACGACTGGGTC	
27	S P I L W L L P L L S V G F A S V L T Q -	
	GTCGGCACCTACATGCTCGCCAAGTACGCCGGGTGCCGGTTCGACCCGACGAGCTCCGGC	
44560	+-----+-----+-----+-----+	44619
	CAGCCGTGGATGTACGAGCGGTTTCATGCGGCCCCGACGGCCAGCTGGGCGTCTCGAGGCCG	
27	V G T Y M L A K Y A G L P V D P Q S S G -	
	GTCCTGATGGTCCTCGTGTTCCGGTGTCGGCACCGACTACGCCCTGCTGCTCATCGCCCGC	
44620	+-----+-----+-----+-----+	44679
	CAGGACTACCAGGAGCACAAAGCCACAGCCGTGGCTGATGCGGGACGACGAGTAGCGGGCG	
27	V L M V L V F G V G T D Y A L L L I A R -	
	TACCGTGAGGAACTGCGCCGCGAGCAGGACCGGCACGTGGCCATGAAGACCGGTTGCGA	
44680	+-----+-----+-----+-----+	44739
	ATGGCACTCCTTGACGCGGGCGCTCGTCTGGCCGTGCACCGGTACTTCTGGCGCAAACGCT	
27	Y R E E L R R E Q D R H V A M K T A L R -	
	CGGTTCGGGGCCCGGCCATCCTGGCCTCGGCCGGCACCATCGCCATCGGCCTCGTCTGCCTG	

[illegible]

CTTCCGCGATCATCCAGATGTCCGCGGTCTGACCGCCCGGAGCCCCGGCCCATCGCGGCG
47140 +-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 47199
GAAGGCGCTAGTAGGTCTACAGGCGCCAGACTGGCGGGCCTCGGGGCCGGGTAGCGCCGC
28 S A I I Q M S A V * (ORF28)

CGGGCCACGGCAGACAAGGAGAGAGCGTATGGCCGGCCTGGTCATGTGCGCCGGTGGAGGC
47200 +-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 47259
GCCCGGTGCCGTCTGTTCTCTCTCGCATACCGGCCGACCAGTACAGCGGCCACCTCCG
(ORF29) M A G L V M S P V E A -

GCTCGACGCGCTGGGCACGGTGCAGGGGCGTCAGGACCCCTATCCCTTCTACGAGGCGAT
47260 +-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 47319
CGAGCTGCGCGACCCGTGCCACGTCCCCGCGAGTCTGGGGATAGGGAAGATGCTCCGCTA
29 L D A L G T V Q G R Q D P Y P F Y E A I -

CCGCGCGCACGGGCAGGCGGTCCCCACGAAGCCCGGCCGCTTCGTGGTGGTTCGGCCACGA
47320 +-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 47379
GGCGCGCGTGGCCGTCCGCCAGGGGTGCTTCGGGCCGGCGAAGCACCACCGCGGTGCT
29 R A H G Q A V P T K P G R F V V V G H D -

CGCGTGCAGCCGGGCGCTGCGGGAACCGGCCCTGCGCGTCCAGGACGCCAGGAGCTACGA
47380 +-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 47439
GCGCACGCTGGCCCGCGACGCCCTTGGCCGGGACGCGCAGGTCTGCGGTCTCGATGCT
29 A C D R A L R E P A L R V Q D A R S Y D -

CGTCGTCTTCCCTCGTGGCGGTGCACTCCTCGGTCCGGGGGTTACCAGCTCCATGCT
47440 +-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 47499
GCAGCAGAAGGGGAGCACCGCCAGCGTGAGGAGCCAGGCCCCCAAGTGGTTCGAGGTACGA
29 V V F P S W R S H S S V R G F T S S M L -

CTACAGCAACCCGCCCCGATCACGGCCGGTTGCGCCAGGTGGTGAGCTTCGCGTTCACCCC
47500 +-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 47559
GATGTCGTTGGGCGGGCTAGTGCCGGCCAACCGGTCCACCACTCGAAGCGCAAGTGGGG
29 Y S N P P D H G R L R Q V V S F A F T P -

GCCCAAGGTGCGCCGGATGCACGGGGTGATCGAGGACATGACCGACCGGCTCCTCGACCG
47560 +-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 47619
CGGGTTCCACGCGGCCTACGTGCCCCACTAGCTCCTGTACTGGCTGGCCGAGGAGCTGGC
29 P K V R R M H G V I E D M T D R L L D R -

GATGGCCCGGCTCGGCTCCGGCGGCTCCCCGGTTCGACCTCATAGCCGAGTTCGCCGCCCG
47620 +-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 47679
CTACCGGGCGGAGCCGAGGCCGCGAGGGGCCAGCTGGAGTATCGGCTCAAGCGGCGGGC
29 M A R L G S G G S P V D L I A E F A A R -

GCTGCCCCGTCGCGGTGATCAGCGAGATGATCGGCTTTCCGGCGAAGGACCAGGTGTGGTT
47680 +-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 47739
CGACGGGCGCGCCACTAGTCGCTCTACTAGCCGAAAGGCCGCTTCTGGTCCACACCAA
29 L P V A V I S E M I G F P A K D Q V W F -

CCGCGACATGGCCTCCCGGGTTCGCCGTGGCGACGGACGGTTTCACCGACCCCGGCGCGCT
47740 +-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 47799
GGCGCTGTACCGGAGGGCCAGCGGCACCGCTGCCTGCCAAAGTGGCTGGGGCCGCGCGA
29 R D M A S R V A V A T D G F T D P G A L -

CACGGGGGCCGACGCCGCCATGGACGAGATGAGCGCCTACTTCGACGACCTCCTGGACCG
47800 +-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 47859
GTGCCCCCGGCTGCGGCGGTACCTGCTCTACTCGCGGATGAAGCTGCTGGAGGACCTGGC
29 T G A D A A M D E M S A Y F D D L L D R -

TCGCCGCCGACCCCCGGCCGACGACCTGGTCACCCTGCTCGCCGAGGCCACGACGGCTC
47860 +-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 47919
AGCGGCGGCGTGGGGCCGGCTGCTGGACAGTGGGACGAGCGGCTCCGGGTGCTGCCGAG
29 R R R T P A D D L V T L L A E A H D G S -

CCCCGGGCGCCTGGACCACGACGAAGTATGGGCACCATGATGGTGCTGCTCACAGCCGG

```

47920 +-----+-----+-----+-----+-----+-----+-----+-----+ 47979
      GGGGCCCCGCGGACCTGGTGCTGCTTGACTACCCGTGGTACTACCACGACGAGTGTTCGGCC
29      P G R L D H D E L M G T M M V L L T A G -

      GTTCGAGACCACGAGCTTTCTGATCGGCCACGGGGCGATGATCGCCCTCGAACAACGGGC
47980 +-----+-----+-----+-----+-----+-----+-----+-----+ 48039
      CAAGCTCTGGTGCTCGAAAGACTAGCCGGTGCCCCGCTACTAGCGGGAGCTTGTTCGGCCG
29      F E T T S F L I G H G A M I A L E Q R A -

      GCACGCGGCCCCGGCTGCGGGCCGAACCCGACTTCGCGGACGGCTACGTTCGAGGAGATCCT
48040 +-----+-----+-----+-----+-----+-----+-----+-----+ 48099
      CGTGCGCCGGGCGGACGCCCCGGCTTGGGCTGAAGCGGCTGCCGATGCAGCTCCTCTAGGA
29      H A A R L R A E P D F A D G Y V E E I L -

      CAGGTTTCGAGCCGCGGTCCACGTCAACGACCCGGTGGGCTGCCGAGGACCTCGACCTGCT
48100 +-----+-----+-----+-----+-----+-----+-----+-----+ 48159
      GTCCAAGCTCGGCGGCCAGGTGCAGTGGTTCGGCCACCCGACGGCTCCTGGAGCTGGACGA
29      R F E P P V H V T S R W A A E D L D L L -

      GGGCCTGTCCGTACCCGCGGGCTCCAAGCTGGTCTGATCCTGGCCGCGCGGAATCGCGA
48160 +-----+-----+-----+-----+-----+-----+-----+-----+ 48219
      CCCGGACAGGCATGGCCGCCCCGAGGTTTCGACCAGGACTAGGACCGGCGGCGCTTAGCGCT
29      G L S V P A G S K L V L I L A A A N R D -

      TCCCGGCCGCTACCCCGAGCCCGGCGCTTCGACCCCGACCGCTACGCGCCCCGGCCGGG
48220 +-----+-----+-----+-----+-----+-----+-----+-----+ 48279
      AGGGCCGGCGATGGGGCTCGGGCCGGCGAAGCTGGGGCTGGCGATGCGCGGGGCGCGCCC
29      P G R Y P E P G R F D P D R Y A P R P G -

      CGGGCCGGAGGCCACCAGACCGCTGAGCTTCGGCGCGGGCGGCCACTTCTGCCTCGGCGC
48280 +-----+-----+-----+-----+-----+-----+-----+-----+ 48339
      GCCCGGCCTCCGGTGGTCTGGCGACTCGAAGCCGCGCCCGCGGTGAAGACGGAGCCGCG
29      G P E A T R P L S F G A G G H F C L G A -

      TCCGCTGGCGCGGCTGGAAGCCCGGATCGCGCTGCCGCGTCTGCTGCGCCGCTTCCCGGA
48340 +-----+-----+-----+-----+-----+-----+-----+-----+ 48399
      AGGCGACCGCGCCGACCTTCGGGCCTAGCGCGACGGCGCAGACGCGCGCGAAGGGCCT
29      P L A R L E A R I A L P R L L R R F P D -

      CCTGGCCGTGTCCGAGCCCCCGTCTACCGCGACCGCTGGGTTCGTCCGCGGCCTCGAAAC
48400 +-----+-----+-----+-----+-----+-----+-----+-----+ 48459
      GGACCGGCACAGGCTCGGGGGGCGAGTGGCGCTGGCGACCCAGCAGGCGCCGGAGCTTTG
29      L A V S E P P V Y R D R W V V R G L E T -

      CTTTCCCGTGACCCTCGGGTCTGAGCCCCCGCGGCCGGAACACGTGACCGTCCCGGCC
48460 +-----+-----+-----+-----+-----+-----+-----+-----+ 48519
      GAAAGGGCACTGGGAGCCCAGGACTCGGGGGCGGCCGCTTGTGCACTGGCAGGGCCGG
29      F P V T L G S * (ORF29)

      GGCGGGTGCGCGCCCTCTCAGACGTACAGGGTGTGGGGCCCTGACCACACAGCACCCGG
48520 +-----+-----+-----+-----+-----+-----+-----+-----+ 48579
      CCGCCACGCGCGGGAGAGTCTGCATGTCCACAAACCCGGGACTGGTGTGTCTGGGCC

      CCGTACAGCTCCAGGTGGTGCTCGGGTTCATGCAGGTGCAGCGTGATGCTCTGGGCATC
48580 +-----+-----+-----+-----+-----+-----+-----+-----+ 48639
      GGCATGTCGAGGTCCAACCACGAGCCCCAGTACGTCCACGTGCACTACGAGACCCGTAG
30      (ORF30)* A P A A H H E P C

      GCTGCACGCGCTGGATCGGGACGTCTGTGTAGATCGAGGACCCGCGCTCGCCTGGGCGA
48640 +-----+-----+-----+-----+-----+-----+-----+-----+ 48699
      CGACGTGCGCGACCTAGCCCTGCAGCAACACTAGCTCCTGGGCGGCGAGCGGACCCGCT
30      R Q V R Q I P V D N Y I S S G G S A Q A -

      GGATGTCCACCGACTCCTTGCCCAGTCGGCACGCCCCGCCCCAGCAGGCCGCGGCACAGCA
48700 +-----+-----+-----+-----+-----+-----+-----+-----+ 48759
      CCTACAGGTGGCTGAGGAACGGGTTCAGCCGTGCGGGCGGGTTCGTCCGGCGCCGTGTCGT

```

30	L I D V S E K G L R C A R G L L G R C L -	
	CCCCCTCCTCCAGCGTCCAGGCCCTCGCCCCGAAGCCCCCTTGAGTGCACGAGGTTCGGCCA	
48760	+-----+-----+-----+-----+-----+-----+-----+	48819
	GGGCGAGGAGGTTCGACGGTCCGGAGCGGGCTTCGGGGGAACCTCAGCTGCTCCAGCCGGT	
30	V R E E L T W A E G S A G K S D V L D A -	
	GCCGATGGGCGTGAAACCGTGCCTCGTCGGCCAGCAGGGTTCGCCTCGCCGAGCTGCAGGT	
48820	+-----+-----+-----+-----+-----+-----+-----+	48879
	CGGCTACCCGCACCTTGGCACGGAGCAGCCGGTTCGTCACGCGGAGCGGCTCGACGTCCA	
30	L R H A H F R A E D A L L T A E G L Q L -	
	GGGTGATCGGCGCCGAGCCCTGCTCCTCGTACTCGGTGTAGGTGATCTTTCGGCCGGGCA	
48880	+-----+-----+-----+-----+-----+-----+-----+	48939
	CCCCTAGCCGCGGCTCGGGACGAGGAGCATGAGCCACATCCACTAGAACCCTCGGCCCGT	
30	H T I P A S G Q E E Y E T Y T I K R G P -	
	GCCTCCCGCGGAAGACGTCTTGAGCGGCCGCGGCCAGTCCGGTCATGGTGCCGACCGACG	
48940	+-----+-----+-----+-----+-----+-----+-----+	48999
	CGGAGGGCGCCTTCTGCAGGACTCGCCGGCGCCGGTCAGGCCAGTACCACGGCTGGCTGC	
30	L R G R F V D Q A A A A L G T M T G V S -	
	AGGCCGAGGCCACGGCCAGCATCGGCGCCCGGAACATCGGTGATCCGGCGTTGAGTTCGG	
49000	+-----+-----+-----+-----+-----+-----+-----+	49059
	TCCGGCTCCGGTGCCTGCTAGCCGCGGGCCTTGTAGCCACTAGGCCGCAACTCAAGCC	
30	S A S A V A L M P A R F M P S G A N L E -	
	AGGCGTACTGCTGCTGGAGCACCGCGCCAGCGGAAGGACGCGCTCCTGGGGAACGAAGA	
49060	+-----+-----+-----+-----+-----+-----+-----+	49119
	TCCGCATGACGACGACCTCGTGGCGCGGGTTCGCTTCTGCGCGAGGACCCCTTGCTTCT	
30	S A Y Q Q Q L V A G L P L V R E Q P V F -	
	CGTCCGCGGCGATGGTGCTGACGCTTCCCAGCCCCGAGCCCCGAGGTGTGCCAGTTCGT	
49120	+-----+-----+-----+-----+-----+-----+-----+	49179
	GCAGGCGCCGCTACCACGACTGCGAAGGGCTCGGGGCTCGGGGCTCCACACGGTCAGCA	
30	V D A A I T S V S G S G R L G S T H W D -	
	CGACGATCTGCAGCTGGTTCGGTCGGCACAGGGCCATCACGGGCTGCATGCCGCCGTCGG	
49180	+-----+-----+-----+-----+-----+-----+-----+	49239
	GCTGCTAGACGTCGACGAGCCAGCCGTTGGTCCCGGTAGTGCCGACGTACGGCGGCAGCC	
30	D V I Q L Q D T P V L A M V P Q M G G D -	
	GGGTTCGGTGAGACGGCGATCAGAACCTGCCAGTGACTGTGCCAGGCACCGCTGATGAAGC	
49240	+-----+-----+-----+-----+-----+-----+-----+	49299
	CCCAGCCACTCTGCCGCTAGTCTTGGACGGTCACTGACACGGTCCGTGGCGACTACTTCG	
30	P T P S V A I L V Q W H S H W A G S I F -	
	CCCCTTGGCGTTCACTACGACACCGCCGTCGACCGGGGCCCATGCCGCCGGGACTGA	
49300	+-----+-----+-----+-----+-----+-----+-----+	49359
	GGGTGAACGGCAAGTGATGCTGTGGCGGCAGCTGGCCCCGGCGGTACGGCGGCCCTGACT	
30	G W K G N V V V G G D V P A A M G G P S -	
	GGGTGCCGAGACCCGGACATCCGGCCGGGAGAACACCTCGTCTGACGTGGTTCGGGGA	
49360	+-----+-----+-----+-----+-----+-----+-----+	49419
	CCCACGGCCTCTGGGCTGTAGGCCGGCCCTCTTGTGGAGCAGGACGTGCACAGCCCCCT	
30	L T G S V R V D P R S F V E D Q V H D P -	
	AGAGGCCCCGCATCCAGGTGGGTATCCACCACACCGAGGCCGTCCAGGCGGCCGATCCGT	
49420	+-----+-----+-----+-----+-----+-----+-----+	49479
	TCTCCGGGCGTAGGTCCACCCATAGGTGGTGTGGCTCCGGCAGGTCCGCGCGCTAGGCA	
30	F L G A M W T P I W W V S A T W A A S G -	
	CGCCGCGCGCCAGCTCGGCGGCCACGTCCACCAGGGTGCGGGCGTCCGACTCGAAGCCGC	
49480	+-----+-----+-----+-----+-----+-----+-----+	49539
	GCGGCGCGCGGTTCGAGCCCGGTGCAGGTGGTCCCACGCCGAGCCTGAGCTTCGGCG	
30	D G R A L E A A V D V L T R A D S E F G -	

[illegible]

33 H A P T M N S P T S P S R C S S P S G P -
 CTTTCGGCGCGCCGGATCCCGCGGAACGGTTCGGGCCGGAGACGGCAGAGCGGTCACTGG
 53800 +-----+-----+-----+-----+-----+-----+-----+----- 53859
 GGAAGCCGCGCGGCCCTAGGGCGCCTTGCCAAGGCCGGCCTCTGCCGTCTCGCCAGTGACC
 33 F G A P D P A E R F R P E T A E R S L V -
 TCACTTTCGCCACCTCCAGGGGCATGTGTTCGGCTGCATCGGCTTCCC GCCACGGTACGGG
 53860 +-----+-----+-----+-----+-----+-----+-----+----- 53919
 AGTGAAGCGGTGGAGGTCCCCGTACACAGCCGACGTAGCCGAAGGGCGGTGCCATGCCC
 33 T F A T S R G M C R L H R L P A T V R E -
 AGCACATGTTGCATGGCAATACCTTTCCAAGTCGGTGGCAACCCCTCCTTGCCATCCACCC
 53920 +-----+-----+-----+-----+-----+-----+-----+----- 53979
 TCGTGTACAACGTACCGTTATGGAAGGTTTCAGCCACCGTTGGGAGGAACGGTAGGTGGG
 33 H M L H G N T F P S R W Q P S L P S T H -
 ACTGCAGTTGGGCGAGATGTGTAGGCATTCGAGGTCCGCAGGTTTGCCAAGCCGCGCGCG
 53980 +-----+-----+-----+-----+-----+-----+-----+----- 54039
 TGACGTCAACCCGCTCTACACATCCGTAAGCTCCAGGCGTCCAAACGGTTTCGGCGCGCGC
 33 C S W A R C V G I R G P Q V C Q A A R D -
 ACCGGCATACTCTCTGGCACAACCTGGAATGAGTAGCGTGGCAGGCCACGGGGACCGGGCC
 54040 +-----+-----+-----+-----+-----+-----+-----+----- 54099
 TGGCCGTATGAGAGACCGTGTGACCTTACTCATCGCACCGTCCGGTGGCCCTGGCCCCG
 33 R H T L W H N W N E * (ORF33)
 GGGCCAGGAACCTTCGTCTCCATCTATTCGCTGGGGCGTGCACGTGTTGGAGCAGCCAT
 54100 +-----+-----+-----+-----+-----+-----+-----+----- 54159
 CCCGGTCTTGAAGCAGGAGGTAGATAAGCGACCCCGCACGTGCACAACCTCGTCGGTA
 CTTTCGGCCGTCGCCCTGAGGCAGCTGAGGACCGAGCGGGTCTTTCCAGGCCGCGCTCG
 54160 +-----+-----+-----+-----+-----+-----+-----+----- 54219
 GAAAGCCGGCAGCGGACTCCGTGACTCCTGGCTCGCCCCAGAAAGGGTCCGGCGCGAGC
 CGGGGGACGGCATGTCTACGGGCTATCTCTCGCGCCTGGAGTCGGGCGCCCGGCAGCCCT
 54220 +-----+-----+-----+-----+-----+-----+-----+----- 54279
 GCCCCCTGCCGTACAGATGCCCGATAGAGAGCGCGACCTCAGCCCGCGGCGCGTCCGGA
 34 (ORF34) M S T G Y L S R L E S G A R Q P S -
 CCGATCGCGCCGTCGCCCACCTGGCCGGACAACCTCGGCATCAGCCCGTCGGAGTTCGAAG
 54280 +-----+-----+-----+-----+-----+-----+-----+----- 54339
 GGCTAGCGCGGCAGCGGGTGGACCGGCCTGTTGAGCCGTAGTCGGGCAGCCCTCAAGCTTC
 34 D R A V A H L A G Q L G I S P S E F E G -
 GGTCCCGGGCCACCTCGCTCGCCAGATCCTCTCCCTCTCCACTTCCCTGGAGTCCGACG
 54340 +-----+-----+-----+-----+-----+-----+-----+----- 54399
 CCAGGGCCCGGTGGAGCGAGCGGGTCTAGGAGAGGGAGAGGTGAAGGGACCTCAGGCTGC
 34 S R A T S L A Q I L S L S T S L E S D E -
 AGACCAGTGAGCTTCTCGCCGAGGCGGTACGTTCCGCGCATGGCCAGGATCCGATGCTCC
 54400 +-----+-----+-----+-----+-----+-----+-----+----- 54459
 TCTGGTCACTCGAAGAGCGGCTCCGCCATGCAAGGCGCGTACCGGTCTTAGGCTACGAGG
 34 T S E L L A E A V R S A H G Q D P M L R -
 GCTGGCAGGCCCTGTGGCTGCTGGGACAGTGGAAGCGCCGGCACGGCGACTCGGCCGGCG
 54460 +-----+-----+-----+-----+-----+-----+-----+----- 54519
 CGACCGTCCGGGACACCGACACCTGTACCTTCGCGGCCGTGCCGTGAGCCGGCCGC
 34 W Q A L W L L G Q W K R R H G D S A G E -
 AGCACGGCTACCTCCAGCGTCTGGTGACGCTGAGTGAGGAGATCGGCCTGGCCGAGTTGC
 54520 +-----+-----+-----+-----+-----+-----+-----+----- 54579
 TCGTGCCGATGGAGGTGCGAGACCACTGCGACTCACTCCTCTAGCCGGACCGGCTCAACG
 34 H G Y L Q R L V T L S E E I G L A E L R -
 GCGCACGGGCCCTGACCCAGTTCGCCCCGTGCTGCGGGTACTGGGCGAGATCGTTCCGG

[illegible]

	57880	+-----+-----+-----+-----+-----+	57939
		GCTCGCGGAGCCGGTCCCCGAAGCCCATAGTCCCAGCCGTGCCAGCGGCACAACCCCGGCG	
37		A L A E A L A E P I L T P V T A T N P G -	
		GCATGCAGGCGATGCGCTGGCGTCCCCGCGCCACCAGGGTCTCGCCGCCGTTCGTGCCCCA	
	57940	+-----+-----+-----+-----+-----+	57999
		CGTAGCTCCGCTACGCGACCCGAGGGGCGCGGTGGTCCCAGAGCGGCGGCAGCAGCGGGT	
37		R M C A I R Q R G R A V L T E G G D D G -	
		GCTTGATGTAGTCTGAAGGTGAACTCCAGCTGGGTCTGCCGCAGCTCCGAGAGCCTCATCC	
	58000	+-----+-----+-----+-----+-----+	58059
		CGAACTACATCAGCTTCCAATTGAGGTGACCCAGACGGCGTTCGAGGCTCTCGGAGTAGG	
37		L K I Y D F T F E L Q T Q R L E S L R M -	
		GGATCGACAGTTTCGTCTGAAGGCGGTGATCTCCGCGAAGAACTCGCAGTCCACCTTGAGGG	
	58060	+-----+-----+-----+-----+-----+	58119
		CCTAGCTGTCAAGCAGCTTCCGCCACTAGAGGCGCTTCTTGAGCGTCAGGTGGAAGTCCC	
37		R I S L E D F A T I E A F F E C D V K L -	
		TGAAGAGCTTGAGGTCTCTCGGACCTCGGCGAGCACCGAAGGCGCCCTCTCCTTGAGAA	
	58120	+-----+-----+-----+-----+-----+	58179
		ACTTCTCGAACTCCAGGAGGACCTGGAGCCGCTCGTGGCTTCCGCGGGAGAGGAACTCTT	
37		T F L K L D E Q V E A L V S P A R E K L -	
		AGAGTTCCCCGCAACGCCCCCTGCCAACGAAGGTAGTTGACGTAGTAGACGTTGCCGACGA	
	58180	+-----+-----+-----+-----+-----+	58239
		TCTCAAGGGCCGTTGCGGGGACGGTTGCTTCCATCAACTGCATCATCTGCAACGGCTGCT	
37		F L E R C R G Q W R L Y N V Y Y V N G V -	
		GGTTCGTCTCCTCGAAGCCGACGGTGTGGCGGAGCTCGAAGTAGTCAGGATTTCGTGCGGG	
	58240	+-----+-----+-----+-----+-----+	58299
		CCAAGCAGAGGAGCTTCGGTGTCCACACCGCCTCGAGCTTCATCAGTCCTAAGCAGCGCC	
37		L N T E E F G V T H R L E F Y D P N T A -	
		TCATAGGTCTGTGCCCTTCGTTCGTGCGGGCCGGTTCGTTCGACCCGAGTTGCGTGAAGCAAC	
	58300	+-----+-----+-----+-----+-----+	58359
		AGTATCCAGACACGGGAAGCAGCAGCCCCGGCCAGCAGCGTGGCTCAACGCACTTCGTTG	
37		T M (ORF37)	
		TCACTGGTTCGCGATGGCCTGCGGGGTTCGGTGGCCCGCGCTCCGGGCGGAGTGCGGGCG	
	58360	+-----+-----+-----+-----+-----+	58419
		AGTGACCAGCGCTACCGGACGCCCCAGCCACCGGGCGCGAGGCCCGCCTCTCACGCCCCG	
		GGGTGCCGGCCGGCGCGGGGTTCAGCCGCGCGCCGACGGCAGCAGGGGAAGAACCTCTCG	
	58420	+-----+-----+-----+-----+-----+	58479
		CCCACGGCCGGCCGCGCCCCAGTCGGCGCGCGGCTGCGTTCCTTCTTGGGAGAGC	
38		(ORF38) * G R A S P L L P L V R E -	
		CGGCCGCTCGTGGAGCCGTTCGGGGGCGGTGCGCCGTAGGTGACGGAGATAACCCGGGCTC	
	58480	+-----+-----+-----+-----+-----+	58539
		GCCGGCGAGCACCTCGGCAGCCCCCGGCCACGCGGCATCCACTGCCTCTATGGGGCCGAG	
38		R G S T S G D P A P A G Y T V S I G R S -	
		TGCGCGGCGCGCACGATCCCCGGCATCGCGCGTTTCGGCGAGCGCCGCGATGGTCATCGCG	
	58540	+-----+-----+-----+-----+-----+	58599
		ACGCGCCGCGCGTGTCTAGGGGCGGTAGCGCGCAAGCCGCTCGCGGCGCTACCACTAGCGC	
38		Q A A R V I G P M A R E A L A A I T M A -	
		GGATTGACCGTCAGCGCGCCGGGAACCGACGATCCGTTCGGTGACGAAGATCCCCGGGTGG	
	58600	+-----+-----+-----+-----+-----+	58659
		CCTAACTGGCAGTCGCGCGGCCCTTGGCTGCTAGGCAGCCACTGCTTCTAGGGGCCCCACC	
38		P N V T L A G P V S S G D T V F I G P H -	
		TCGCGGAGCTCGTTGCTGTCTCGTCCAGGGCGGATGTGTGGGGGTTCGTGCCCCATCCGGCAG	
	58660	+-----+-----+-----+-----+-----+	58719
		AGCGCCTCGAGCAACGACAGCAGGTCCCGCTACACACCCCCAGCAGCGGGTAGGCCGTC	

CGCGCCCCATGGGCGATGGCGGCGGGCAGGTAGTTGACCAGCAAGGACTGCTTGGCCTCG

```
59500 +-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 59559
GCGCGGGGTACCCGCTACCGCCGCGCCGTCATCAACTGGTTCCTGACGAACCGGAGC
38   R A G H A I A A P L Y N V L L S Q K A E -

AAGCGGCAGCCGCGCCATCATCCAGTTGCAGTTCACGCACTTGGTGTGTGTCGATGGCGACG
59560 +-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 59619
TTCGCCGTGCGCCGGTAGTAGGTCAACGTCAAGTGCCTGAACCACAACAGCTACCGCTGC
38   F R C G A M M W N C N V C K T N D I A V -

GCGAGGGGGTTGGCGGTGCGGCCGCGTGGTTGCACGCCGCGGCCACAGTCCGCCGGCG
59620 +-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 59679
CGCTCCCCCAACCGCCACGCCGCGCCACCAACGTGCGGCGCGGGGTGTCAGGCGGCCGC
38   A L P N A T R G A H N C A A A W L G G A -

TAGCTCACGTCGTTCCAGTCTCTGCCGGTTCACGGAGAGGGACTCCTCGACACGGTCGTAC
59680 +-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 59739
ATCGAGTGCAGCAAGGTTCAGGACGCCAGTGCCTCTCCCTGAGGAGCTGTGCCAGCATG
38   Y S V D N W D Q R T V S L S E E V R D Y -

CAGGGGTCCAGGGTTTCGCGGCTCACCGCCTGCGGCCACATCCGGCGTCCTATGGACCCC
59740 +-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 59799
GTCCCCAGGTCCCAAAGCGCCGAGTGGCGGACGCCGGTGTAGGCCGAGGATACCTGGGG
38   W P D L T E R S V A Q P W M R R G I S G -

TGCCGGTTCGAAGACGAAGCGCGGGGCGCGGGGCATCGCGGCGAAGTAGACGACGCTGCCG
59800 +-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 59859
ACGGCCAGCTTCTGCTTCGCGCCCCGCGCCCCGTAGCGCCGCTTCATCTGCTGCGACGGC
38   Q R D F V F R P A R P M A A F Y V V S G -

CCGCCCACACAGTTCCTCCGCGAGGATGCTCATGCCGTCCCCGACCGTGAAGTCGAACGCC
59860 +-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 59919
GGCGGGTGTGTCAAGGGCGGCTCCTACGAGTACGGCAGGGGCTGGCACTTCAGCTTGCGG
38   G G V C N G G L I S M G D G V T F D F A -

CTCGTGACGAGGAGCCGAGTTTGTAGTCGTGCTCGAACTCCTTGCTCTCCAGCCACGGC
59920 +-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 59979
GAGCACATGCTCCTCGGCTCAAACATCAGCACGAGCTTGAGGAACGAGAGGTCCGTGCCG
38   R T Y S S G L K Y D H E F E K S E L W P -

CCGCGTTCCAGGACGGTGACGTCCGCGCCCCCGCCGCGCCAGGTGGTAGGCGCGCATGGCA
59980 +-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 60039
GGCGCAAGGTCCCTGCCACTGCAGCCGCGGGGGGCGGCGGTCCACCATCCGCCGCTACCGT
38   G R E L V T V D A G G A A L H Y A A I A -

CCGCCGAATCCGCTGCCGATGACGAGGACGTCCGTGCGCTCGGCCGTGGTGCTCATGCGG
60040 +-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 60099
GGCGGCTTAGGCGACGGCTACTGCTCCTGCAGGCACGCGAGCCGGCACCACGAGTACGCC
38   G G F G S G I V L V D T R E A T T S M
                                     (ORF39) * A -

GGCTCCCGGTGGACGTGGTGTGCGGGTGGAGGCGGGCGAACTCACGCCCGTAGCTGTAAT
60100 +-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 60159
CCGAGGGCCACCTGCACCACAGCCACCTCCGCCCGCTTGAGTGCGGGCATCGACATTA
39   P S G T S T T D P H L R A F E R G Y S Y -

CCTTGAAGCGCCACAGGCCGTGCGCGTCCGGCATGCTCAGGCCCATGGCCTCCAGTCCCCG
60160 +-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 60219
GGAACTTCGCGGTGTCCGGCAGCCGAGCCGTACGAGTCCGGGTACCGGAGGTGAGGGC
39   D K F R W L G D A D P M S L G M A E L G -

GATGGCCGTCTCCATCGCCTGTGCCGTGTTGAGGTGCGCGGCCGAATCGAAGGCCATGT
60220 +-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 60279
CTACCGGCAGGAGGTAGCGGACACGGCACAACCTCCACGCGCCGGCTTAGCTTCCGGTACA
39   P H G D E M A Q A T N L H A A S D F A M -
```


61900 GAGCTGGCGCAGCACTGCTCGGACGCCTGGCACCGGTTTCGCCGAGGAGATGGTTCGGCGC 61959
 +-----+-----+-----+-----+-----+-----+-----+-----+
 40 CTGACCGCGTCGTGACGAGCCTGCGGACCGTGGCCAAGCGGCTCCTCTACCAAGCCGCG
 E L A Q H C S D A W H R F A E E M V R R -

 61960 CGGCACAGCGCCGCGGCCGACGGCTCCGGACTCCGCACGACTCAGCCAACCTCGCCGTTGA 62019
 +-----+-----+-----+-----+-----+-----+-----+-----+
 40 GCCGTGTGCGCGCGCCGCTGCCGAGGCCTGAGGCGTGTGAGTCGGTTGAGCGGCAACT
 R H S A A A D G S G L R T T Q P T R R *
 (ORF40) -

 62020 CGGCCTACTTCGACAGGGAGTTACGGTGACCACGAACACCATCGAGGACGCGGTCCGCCG 62079
 +-----+-----+-----+-----+-----+-----+-----+-----+
 41 GCCGGATGAAGCTGTCCCTCAATGCCACTGGTGCTTGTGGTAGCTCCTGCGCCAGGCGGC
 (ORF41) V T T N T I E D A V R R

 62080 GGTCGTGAGTACATGCACGTCAACCTGGGTGAGAACCTCACGATCGATGACATGGCGCG 62139
 +-----+-----+-----+-----+-----+-----+-----+-----+
 41 CCAGCAGCTCATGTACGTGCACTGGACCCAGTCTTGGAGTGCTAGCTACTGTACCGCGC
 V V E Y M H V N L G Q N L T I D D M A R

 62140 CACGGCGATGTTTCAGCAAGTTCATTTACCCCGCATCTTCCGCGAAGTCACCGGTACCTC 62199
 +-----+-----+-----+-----+-----+-----+-----+-----+
 41 GTGCCGCTACAAGTCGTTCAAGGTAAAGTGGGCGTAGAAGGCGCTTCAGTGGCCATGGAG
 T A M F S K F H F T R I F R E V T G T S -

 62200 TCCCGGGCGTTTCTGTCCGCCTTACGGATTTCAGGAGGCCAAGAGACTTCTCGTGCACAC 62259
 +-----+-----+-----+-----+-----+-----+-----+-----+
 41 AGGGCCCGCAAAGGACAGGCGGAATGCCTAAGTCTCCGTTCTCTGAAGAGCACGTGTG
 P G R F L S A L R I Q E A K R L L V H T -

 62260 TGCCTCAGTGTGGCCGATATCAGCAGTCAGGTCCGCTACAGCAGTGTGGTACTTTTCAG 62319
 +-----+-----+-----+-----+-----+-----+-----+-----+
 41 ACGTGAGTCACACCGGCTATAGTCAGTCAGTCCAGCCGATGTGTCACAGCCATGAAAGTC
 A L S V A D I S S Q V G Y S S V G T F S -

 62320 TTCTCGCTTCAAGGCCTGTGTGGGGCTTTCCCCGAGCGCCTATCGCGACTTCGGCGGGGT 62379
 +-----+-----+-----+-----+-----+-----+-----+-----+
 41 AAGAGCGAAGTTCCGGACACACCCCGAAAGGGGCTCGCGGATAGCGCTGAAGCCGCCCA
 S R F K A C V G L S P S A Y R D F G G V -

 62380 GCAGCCGGGTTTTCCCTCCGCCGCGGCCCGTCTCACTCCACCGCGCACAAATCCCTCCGT 62439
 +-----+-----+-----+-----+-----+-----+-----+-----+
 41 CGTCCGCCCCAAAAGGGAGGCGGCGCGGCGAGAGTGGGCGTGTAGGGAGGCA
 Q P G F P S A A A R L T P T A H N P S V -

 62440 GCGCGGCCGCATTCACTCCGCCCGGGTGACAGGCCCGGAAGGATCTTCGTGGGCCTGTT 62499
 +-----+-----+-----+-----+-----+-----+-----+-----+
 41 CGCGCCGGCGTAAGTGAGGCGGGGCCACTGTCCGGGCCTTCTAGAAGCACCCGGACAA
 R G R I H S A P G D R P G R I F V G L F -

 62500 CCCCCGAGGATGCGCCAGGGCCCGCGCGCTGGACCGTCATGGAGAGTCCCGGGGC 62559
 +-----+-----+-----+-----+-----+-----+-----+-----+
 41 GGGGCGCTCCTACGCGTCCCGGCGGGCCGCGGACCTGGCAGTACCTCTCAGGGCCCCG
 P G R M R Q G R P A R W T V M E S P G A -

 62560 CTTTCGAGTCCGGGACGTGCCCCGTGGGCACCTGGCACATCCTGGTCCACTCCTTCCCCGC 62619
 +-----+-----+-----+-----+-----+-----+-----+-----+
 41 GAAGCTCGAGGCCCTGCACGGGCACCCGTGGACCGTGTAGGACCAGGTGAGGAAGGGGCG
 F E L R D V P V G T W H I L V H S F P A -

 62620 CGGACACCGGCCGCACCAGCTCGACTCCGAACCGCTGTTGCTCGGGCACAGCGGACCGCT 62679
 +-----+-----+-----+-----+-----+-----+-----+-----+
 41 GCCTGTGGCCGGCGTGGTTCGAGCTGAGGCTTGGCGACAACGAGCCCGTGTGCGCTGGCGA
 G H R P H Q L D S E P L L L G H S G P L -

CGTGGTGACCCCGGTGCCCCTGCTCCGGGCCGGCGGCAGCATCCTCTGCGCGCGGTGGACGC
62680 +-----+-----+-----+-----+-----+-----+-----+----- 62739
GCACCACGTGGGGCCACGGGACGAGGCGCGCCGCTGTAGGAGGACGCGCGCCACCTGCG
V V H P G A L L R P A D I L L R A V D A -

CCTCGATCCACCGGTCTGTGGCCCACTTCGCGCTGGAGAGCCGCTCACCTCGCCGTA
62740 +-----+-----+-----+-----+-----+-----+-----+----- 62799
GGAGCTAGGTGGCCAGGACGACCGGGTGAAGCGCGACTCTCGGCGGAGTTGGAGCGGCAT
L D P P V L L A H F A L E S R L T S P Y -
(ORF42) * R A T -

CTCACCGTCATCGGTAGCCCTCCGCGCATCCGCAGGGAGAGCATGGGTTTCGGCAACCGCC
62800 +-----+-----+-----+-----+-----+-----+-----+----- 62859
GAGTGGCAGTAGCCATCGGGAGGCGCGTAGGCGTCCCTCTCGTACCCAAGCCGTTGGCGG
S P S S V A L R A S A G R A W V R Q P P -
S V T M P L G G R M R L S L M P E A V A -

CGGTGTCCGGCGACGGTACGCAGATCGAGATCGCGGGTGACCAGGGCCGTGACGAACACC
62860 +-----+-----+-----+-----+-----+-----+-----+----- 62919
GCCACAGGCGCTGCCATGCGTCTAGCTCTAGCGCCCACTGGTCCCGGCACTGCTTGTGG
G V R R R Y A D R D R G * (ORF41)
R H G A V T R L D L D R T V L A T V F V -

GCCTCCATCATCCCGAGGTTGCTGCCGACGCAGAACCGGGCCCCGCGCCGAACGGGATG
62920 +-----+-----+-----+-----+-----+-----+-----+----- 62979
CGGAGGTAGTAGGGCTCCAACGACGGCTGCGTCTTGGCCCCGGGGCGCGGCTTGCCCTAC
A E M M G L N S G V C F R P G A G F P I -

TACGCGTACCGCGGCCGGTCGGCGGTCTGCCGGGGTTTCAACCGCTCGGGGTCTGAAGCGC
62980 +-----+-----+-----+-----+-----+-----+-----+----- 63039
ATGCGCATGGGCGCCGGCCAGCCGACGACGCGCCCAAGCTTGGCGAGCCCCAGCTTTCGCG
Y A Y R P R D A T Q R P E F R E P D F R -

TCGGGGTCTCTCCACAGCCCCGGATGGCGGTGCATGATGTACGGGCAGACCAGCACATCC
63040 +-----+-----+-----+-----+-----+-----+-----+----- 63099
AGCCCCAGGAGGGTGTTCGGGGCCTACCGCCACGTACTACATGCCCGTCTGGTCTGTAGG
E P D E W L G P H R H M I Y P C V L V D -

GATCCGGCGGACACCGTGTAGCCGCCGACCACATCGCGTTGCTGGGCCACCCTGGGCAGG
63100 +-----+-----+-----+-----+-----+-----+-----+----- 63159
CTAGGCCCGCTGTGGCACATCGGCGGCTGGTGTAGCGCAACGACCCGTTGGGACCCGTTCC
S G A S V T Y G G V V D R Q Q A V R P L -

ATCCC
63160 +---- 63164
TAGGG
I G -